Author:
Neville, Vikki M

Title:
Affect and Decision Making

Disentangling Underlying Processes

Affect and Decision Making; Disentangling Underlying Processes

Vikki Maria Neville

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of in the Faculty of Health Sciences.

Bristol Veterinary School
September 2019

53,571 words

Abstract

Judgement bias, which examines decision-making under ambiguity, is a promising measure of affect in non-human animals. However, there are significant gaps in our understanding of the cognitive processes underlying judgement bias and the adaptive significance of affect-induced shifts in decision-making, which we aimed to resolve. To this end, we investigated the relationship between reward and punisher experience, judgement bias, and affective state in rats and humans. Firstly, we conducted a systematic review and meta-analysis of pharmacological manipulations of judgement bias which demonstrated that pharmacological manipulations designed to alter affect overall influenced judgement bias in the predicted direction, hence supporting the validity of judgement bias as a measure of affect (Chapter 2). Then, we combined behavioural experiments with computational modelling, which permitted a deeper insight into the decision-making processes underlying behaviour. Specifically, we examined the influence of fluctuating primary (Chapter 3) and secondary (Chapter 4) rewards and punishers on judgement bias and self-reported affect in humans. Additionally, we manipulated rats' reward and punisher experience either prior to (Chapter 5) or within (Chapter 6) a judgement bias test session. Data from the rat and human judgement bias studies using a within-test manipulation (Chapters 3, 4, and 6) were analysed using a novel computational model in which we defined decision-making as a partially observable Markov decision process. We found that reward and punisher experience, particularly the predictability (in humans) and absolute favourability (in rats) of these experiences, altered judgement bias primarily by influencing reward or punisher sensitivity. Humans that reported more positively valenced affect exhibited greater predictability-dependent modulation of judgement bias. Finally, we assessed the extent to which the explore/exploit trade-off could provide a novel measure of affect and found that inducing a putatively negative affective state increased the rats' tendency to explore, rather than exploit a food source (Chapter 7).

"Wee, sleekit, cowrin, tim'rous beastie,
O, what a pannic's in thy breastie!
Thou need na start awa sae hasty,
Wi' bickering brattle!
I wad be laith to rin an' chase thee,
Wi' murd'ring pattle!"

– Robert Burns, *To a Mouse*, 1785.

Acknowledgements

If you had told 8-year-old me, or even 18-year-old me, that I would have the opportunity to do a PhD in such fascinating fields under the supervision of four world-renowned experts, I wouldn't have believed you. While this opportunity has largely resulted from a lot of chance, and quite frankly a lot of hard work, it would have been unattainable without the help of several people to whom I am completely indebted.

First and foremost, I would like to thank my incredible supervisors: Mike Mendl, Iain Gilchrist, Liz Paul and Peter Dayan for their patience (especially Peter!), wisdom, and guidance over the last four years. I couldn't have asked for a more supportive and inspiring set of supervisors. I would like to especially thank Mike and Liz with whom I have worked most closely. They have both been hugely enthusiastic and knowledgeable, and have never made me feel stupid (even if some of my ideas have been less than great). I have thoroughly enjoyed learning from them and can only hope that I will be even half the supervisor they have been someday. Mike's effort to keep me employed and in academia have also been very welcome, and the final few months of my PhD have been more relaxing without the impending threat of unemployment.

My PhD funding from the BBSRC SWBio-DTP has allowed me to stave off homelessness and starvation, attend some amazing conferences, and conduct several experiments with equipment that was both useful (thanks WPI) and useless (thanks Coulbourn).

I would like to thank my colleagues at Bristol Vet School (especially my officemates Emma, Francesca, and Marco) for providing a very welcoming atmosphere, vegan cake, and for taking part in my (often strange) experiments without hesitation.

To all the rats I have worked with, even the ones who wouldn't co-operate, thank you for being excellent company and for making my many hours stuck a red-lit room more tolerable. Most of all, thank you to the many rats who came to live with me in their retirement: Fattus, Triff, Hookus, Tinky, Mickey, Pooh, Thor, Pebbles, Candycane, Beetle, Newt, G-Unit, Dragon, Jellybean and Crabby. These boys have taught me so much about the weird and wonderful way of rats, provided much entertainment, and have fuelled an incredible love of rats and passion for their welfare. I am also grateful for my two companion rabbits, Daisy and Dandy, who were with me at the start but sadly not the end of my PhD. They furthered my interest animal welfare and tolerated a great deal of academia-related relocation.

There are several people outside of the University of Bristol whom I must acknowledge. I am thankful for my Father (Andy) who has always answered my 'what ifs' with 'let's find out' and has provided an endless supply of 'Dad jokes' and 'facts of the day'. He has been very supportive, interested in my research, and demonstrably proud. I

am also grateful for my one and only younger sibling (Nicole) who provided the practical experience that nicely accompanied the Computational Psychiatry course and who is an incredibly kind, caring, and hilarious individual.

I could write a whole essay on the support that my wonderful partner (Robert) has provided prior to and throughout my PhD, but will refrain from this for the sake of anyone who has decided to read the acknowledgements. I am grateful for the weekends at the beach he sacrificed because I needed to stay at home and work, the meals he cooked when I worked late, his enthusiasm and passion for animal welfare, being an excellent guardian to our many rat boys, his excellent proof-reading of my thesis (sorry about all the errant commas!), and finally for being a shoulder to cry on.

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Signed:

Date:

Contents

List of Figures

List of Tables

Declaration of Involvement

Chapter 2: A version of this chapter is under review at Neuroscience and Biobehavioural Reviews and has been published as a pre-print on BioRxiv:

Neville V, Nakagawa S, Zidar J, Paul ES, Lagisz M, Bateson M, Løvlie H, Mendl M. Pharmacological manipulations of judgement bias: a systematic review and meta-analysis. BioRxiv. 2019 Jan 1:612382.

The division of labour between co-authors is as follows:

Hanne Lovlie, Liz Paul, Melissa Bateson, Mike Mendl and Shinichi Nakagawa conceived the idea. I conducted the literature search and extracted the data jointly with Josefina Zidar and Losia Lagisz. I wrote the R code for effect size calculations and data analysis with input from Shinichi Nakagawa, analysed the data with input from Shinichi Nakagawa and Mike Mendl, and produced the figures (except for Fig. 2.1 which was produced by Losia Lagisz). I wrote the first draft of the manuscript. All authors commented on the draft of this manuscript prior to submission, and three reviewers also provided comments on the manuscript.

Chapter 5: I would like to acknowledge the assistance of Jessica Bonney in the training of rats and collection of data for the progressive ratio lever pressing task described in this chapter, which she undertook as part of her MSc research which was project supervised by myself and Mike Mendl. All other work within this chapter is my own.

Chapter 1

General Introduction

In this general introduction we provide the reader with an overview of both theoretical and empirical research on affect, decision-making, and the relationship between the two. We will start by defining affect, discussing its potential function and underlying neural mechanisms, and describing the importance of measuring affect and current methods available to do this. Then, we will focus on the specific relationship between affect and decision-making, discussing the cognitive mechanisms underlying decision-making and how these processes may be influenced by affect, before outlining judgement bias (a concept central to this thesis) as a measure of affect. Subsequently, we will introduce computational modelling, and discuss its value as a tool to better understand the relationship between affect and decision-making. Finally, we will discuss the aims of this thesis in the context of outstanding questions. To aid the reader and for clarity, a glossary of key terms used in this thesis can be found below (Table 1.1).

1.1 Affective states: an ethological perspective

1.1.1 What is affect?

Defining affect is imperative to its scientific investigation. We start by differentiating between the terms 'emotion', 'mood', and 'affect'. Dictionary definitions for these terms are vague and barely discriminable. For example, according to the Oxford English Dictionary (2019) emotion is 'a strong feeling deriving from one's circumstances, mood, or relationships with others', mood is 'the way you are feeling at a particular time', and affect is defined as 'emotion or desire as influencing behaviour'. Here, we define emotion as referring to transient states, mood as referring to longer term states that persist in the absence of stimuli that would elicit an emotional response, and affect being an umbrella term that encompasses both emotion and mood (Mendl et al., 2010). Moreover, dictionary definitions are too anthropocentric for our purposes. The word 'feeling' implies that affect comprises a subjective component which we cannot say with certainty applies to non-human animals [1]. It is clear from

[1] Although it should be noted that there is consensus in the scientific community that non-human animals may subjectively experience affect; the Cambridge Declaration on consciousness signed by a number of scientists states that 'the weight of evidence indicates that humans are not unique in possessing the neurological substrates that generate consciousness'.

TABLE 1.1: Glossary of key terms used in this thesis

Term	**Definition**
Affect	A multicomponential state, comprising physiology, behaviour, cognition and potentially subjective feelings that is valenced and also varies in terms of arousal.
Affective arousal	The extent to which an affective state promotes behavioural activation or inhibition.
Affective disorders	A pathological condition characterised by abnormalities of affective state.
Affective valence	The extent to which an affective state is pleasant or unpleasant.
Anxiety	An affective disorder (in humans) characterised by a chronically negatively valenced and high arousal affective state.
Bayesian inference	Updating of a belief following new evidence following Bayes' rule (i.e. $P(A\|B) = \frac{P(B\|A)P(A)}{P(B)}$).
Depression	An affective disorder (in humans) characterised by a chronically negatively valenced and low arousal affective state.
Emotion	A short term affective state; typically elicited by rewards or punishers.
Feeling	The subjective component of affect.
Judgement bias	The tendency to opt for either a 'risky' or 'safe' option, particularly when there is ambiguity about the outcome of the 'risky' option.
Mood	A longer term affective state that persists in the absence of rewards and punishers.
Optimism	Opting for a 'risky' as opposed to 'safe' choice more often (or do so with greater alacrity) than others individuals completing the same judgement bias test.
Partially-observable Markov decision process (POMDP)	A framework for defining a reinforcement learning problem where there is uncertainty about the true state of the environment.
Pessimism	Opting for a 'risky' as opposed to 'safe' choice less often (or with weaker alacrity) than others completing the same judgement bias test.
Prior belief	The probability distibution of an outcome before evidence is considered.
Punisher	An outcome that an individual will actively avoid.
Reinforcement learning	A framework for understanding how individuals act to maximise reward/minimise punishment through learning from interactions with their environment.
Reward	An outcome that an individual will work to obtain.
Risk-averse	A tendency to make 'safe' (outcome known) as opposed to 'risky' (outcome unknown) choices.
Risk-seeking	A tendency to make 'risky' (outcome unknown) as opposed to 'safe' (outcome known) choices.
Stress	An acute state of negative valence and high arousal.

the research literature, and indeed our experience as humans, that affect consists of more than a subjective component. James (1884) and Lange (1885) considered the physiological response to be at the core of emotion. They suggested that peripheral feedback precedes feelings; that 'we feel frightened because we are running away'. Other researchers have acknowledged the behavioural component of affect, such as facial expressions (Berridge, 1996; Berridge and Robinson, 2003; Sandem et al., 2002) and vocalisations (Knutson et al., 2002; Watts and Stookey, 2000; Zimmerman and Koene, 1998). Indeed, Darwin (1872) describes the bodily expressions associated with discrete emotions in both human and non-human animals in his seminal book 'The Expression of the Emotions in Man and Animals'. It has also been purported that cognition is an integral part of affect (Paul et al., 2005), with affect suggested to influence a number of cognitive processes including attention (Crump et al., 2018; Mathews and MacLeod, 1994) and memory (Cahill and McGaugh, 1998; Christianson, 1992). According to Lazurus' cognitive-mediational theory (Lazarus, 1991) and appraisal theory (Scherer et al., 2001), emotion arises from our appraisal of a situation; if we see a snake we may not feel scared if we assess that the snake poses no danger in a particular situation.

Dolan (2002) accounted for both the cognitive and physiological components when defining affect as 'psychological or physiological states that index occurrences of value'. Anderson and Adolph's (2014) definition of emotion as 'an internal CNS (central nervous system) state that gives rise to physiological, behaviour, cognitive, and (subjective) responses', encapsulates the multicomponent nature of emotion. This is similar to Roseman's (2008) definition which states that emotion comprises phenomenology (thought and feeling qualities), physiology (neural, chemical, and other physical responses in the brain and body), expressions (signs of emotion state), behaviours (action tendencies or readiness), and emotivations (characteristic goals that people want to attain when the emotion is experienced).

Yet, while these definitions outline the components of emotion, which is undeniably valuable for considering affect in non-human animals, the above definitions cannot be used to distinguish between different affective states. There are two main approaches to characterising affective states: a discrete approach in which affect is described as distinct categories, and a dimensional approach in which affect is defined along a few (usually two or three) dimensions. Ekman (1984) proposed the following categories of emotion: happiness, fear, anger, surprise, grief and sadness. These categories were derived from the observation that these emotions are associated with distinct facial expressions which can be accurately categorised cross-culturally, thought to reflect that distinct and innate biological processes underlie each emotion. Furthermore, Darwin (1872) noted similarities between human and non-human animal expressions in response to threat. Likewise, Panksepp advocated the existence of distinct underlying neurosystems that mediated seeking, rage, fear, lust, care, panic, and play (Panksepp, 1982, 1998). However, meta-analyses of neuroimaging studies have highlighted inconsistencies in the neurocircuitry associated with these discrete emotions (Hamann,

2012; Lindquist et al., 2012).

Further challenges to a discrete approach to defining affect include the existence of culture-specific emotions, such as 'age-otari' in Japan, referring to the feeling of having a bad haircut (Mendl and Paul, 2017). Additionally, by definition, discrete emotions should be entirely distinct. Yet, emotions can be grouped by elemental features; for example, happiness and surprise are more alike than happiness and sadness.

Using factor analysis of subjective reports of emotional words, research has revealed that self-reported affect forms a circumplex structure (Russell, 1980; Russell et al., 1989a; Watson et al., 1999; see Fig. 1.1) supporting a dimensional approach to classifying affect. Valence (the pleasantness or unpleasantness of the affective state), which according to many researchers is the defining characteristic of affect, and arousal (the level of activation) are the most frequently used dimensions (Barrett and Russell, 1999; Carver, 2001; Mendl et al., 2010; Russell, 2003; Watson et al., 1999). For example, excitement is a positively valenced high arousal state, while depression is a negatively valenced low arousal state. Neural substrates have also been proposed to underlie these dimensions. In particular, the mesolimbic system has been associated with affective valence, while the reticular formation has been associated with affective arousal (Posner et al., 2005). Moreover, these dimensions have also been considered in terms of the actions they likely promote and the neurobehavioural systems underlying these actions. Positively valenced high arousal states have been considered to encourage reward acquisition when rewards are abundant, while negatively valenced low arousal states have been considered to promote withdrawal in the absence of obtainable rewards. In contrast, negatively valenced high arousal states been proposed to aid punisher avoidance in high predation environments, while positively valenced low arousal states have been considered to promote relaxation in the absence of punishers to avoid (Boureau and Dayan, 2011; Mendl et al., 2010; see Fig. 1.1). Thus, these dimensions can also be considered in terms of rewards and punishers (relating to valence), and behavioural inhibition and activation (relating to arousal); dopamine been proposed to mediate invigoration in the face of reward and serotonin has been proposed to mediate inhibition when faced with a punisher (Boureau and Dayan, 2011; Mendl et al., 2010). The theory of constructed emotion offers an explanation for the perception of affect as discrete; it suggests that this results from interoception, perception of the situation, and past experience, rather than biologically distinct processes (Barrett, 2014, 2017).

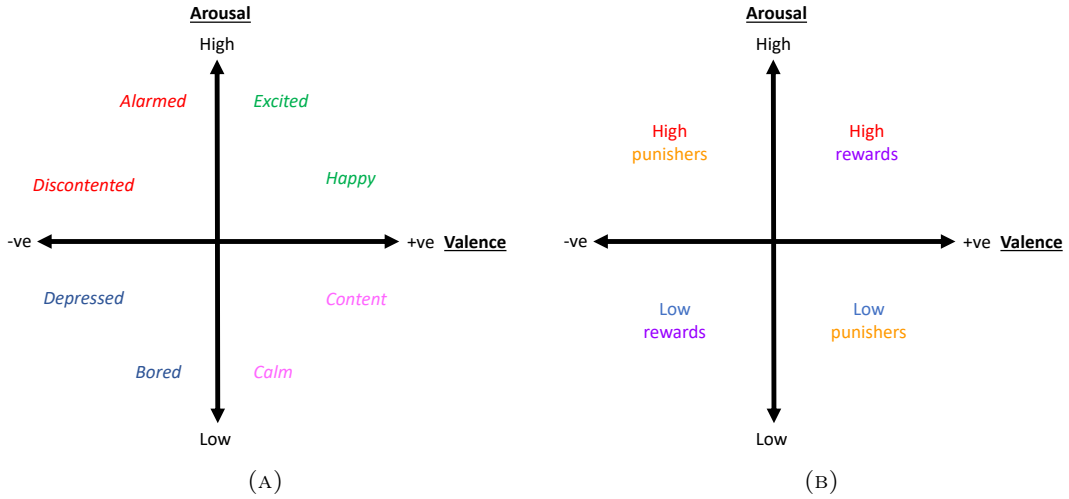


FIGURE 1.1: A dimensional approach to defining affect; (A) a circumplex model of emotion (adapted from Posner et al., 2005) - discrete emotions can be defined according to their position on an arousal axis and a valence axis; (B) an individual's location in this space may reflect their cumulative reward and punisher experience in their environment.

We consider a dimensional approach to be most appropriate for the study of affect in human and non-human animals as it allows the characterisation of affective states without being dependent on human constructs of emotion, thereby obviating the possibility of anthropomorphism. Additionally, it allows us to focus specifically on affective valence and affective arousal, the former being highly relevant to animal welfare, with good welfare being associated with positively valenced affective states and absence of negatively valenced affected states.

To be explicit, we define affect as a multicomponential state, comprising physiology, behaviour, cognitive and potentially subjective feelings that is valenced and also varies in terms of arousal. Additionally, we find Roll's (2013) definition of emotion as 'states elicited by rewards and punishers' helpful as it allows operationalisation of the concept of emotion in non-human animals. Indeed, reward and punisher experience has been proposed to be central to mood and allow adaptive behaviour towards potential future rewards and punishers.

1.1.2 The adaptive significance of affect

It has been suggested that affect functions to allow adaptive behaviour towards fitness-enhancing rewards (e.g. food, water, mates, shelter) and fitness-costly punishers (e.g. predatory attack, thermal damage, poison; Mendl et al., 2010; Nettle and Bateson, 2012; Trimmer et al., 2013). More specifically, affect (particularly mood) might function to coordinate an appropriate response to potential rewards or punishers or a lack thereof given past experience of rewards and punishers, allowing rewards to be maximised and punishers to be minimised. The hypothesis that affect organises and coordinates behaviour is pervasive in the literature, with affect having been suggested

to prioritise goals (Clore et al., 1994; Oatley and Jenkins, 1992) and influence decision-making (Bechara and Damasio, 2005; Dunning et al., 2017; Mendl et al., 2010; Nettle and Bateson, 2012). Similarly, pleasure has been described as providing a common currency for decision-making (Cabanac, 1992).

The somatic marker hypothesis (SMH) states that affect guides decision-making (Bechara and Damasio, 2005; Damasio, 1994). Specifically, the SMH proposes that physiological responses to potential rewarding and punishing outcomes (somatic markers) facilitate decision-making. This has most frequently been studied in the context of risky decision-making, and the finding that individuals who show an impaired physiological response due to brain damage perform more poorly than controls on risk-based behavioural tasks (e.g. the Iowa gambling task) has provided support for the SMH (Bechara and Damasio, 2005; Bechara et al., 1999). However, the SMH has been subject to considerable criticism (Dunn et al., 2006; Maia and McClelland, 2004, 2005), and has been weakened by the finding that control participants have conscious knowledge of the advantageous strategy, meaning that somatic markers may not necessarily inform decision-making (Maia and McClelland, 2004).

Signal detection theory (Green et al., 1966) has also been used to explain the adaptive significance of affective states. Sensory information can be noisy and there may be some degree of overlap between cues that signal rewards or punishers and background noise. To illustrate this, consider the example of a prey animal who has just heard a rustling noise which could either signal an approaching predator or could just be the sound of wind moving vegetation. The individual must decide whether to flee, which would be energetically costly but avoid the risk of predation, or ignore the noise and conserve energy but risk predation. An anxious mood resulting from, for example, repeated encounters with predators or a weakened ability to evade predators due to poor health might result in a reduced threshold for responding to potentially threatening signals (Bateson et al., 2011). However, an individual who is relaxed, reflecting their good health and few punishing experiences, might have a higher response threshold for potential threats (Bateson et al., 2011; Nettle and Bateson, 2012). Similarly, affective state might modulate an individual's threshold for responding to potential rewards. Additionally, high arousal affective states which promote activity, such as anxiety or excitement, may be beneficial when there are rewards to be approached, or punishers to be avoided, while low arousal affective states might allow the conservation of energy when there are few punishers to evade or rewards to be obtained. Thus, mood has been proposed as a system which encodes and integrates information about the state of the individual and their environment to optimise responses to signals indicating potential fitness-enhancing rewards and potential fitness-costly punishers (Bateson et al., 2011; Mendl et al., 2010; Nettle and Bateson, 2012; Trimmer et al., 2013). A core assumption of this hypothesis is that affect is a state which is driven by experience of reward and punisher, and success at acquiring or avoiding them, which will be investigated in Chapters 3, 4, 5, and 6.

1.1.3 Measuring affect in non-human animals

Typically, human affect is measured through self-report (e.g. Anderson et al., 2012; Iigaya et al., 2016; Paul et al., 2011). Numerous questionnaires have been developed for this purpose, such as the affect grid (Russell et al., 1989b), the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988), revised Life Orientation Test (LOT-R; Scheier et al., 1994), and Beck Depression Inventory (BDI; Beck et al., 1961). However, the accuracy of these measures depends on the willingness of a person to disclose their affective state; for example, an individual may conceal their mood to avoid judgement or to appease the experimenter and it also depends on the individual's ability to accurately assess their own affective state. However, as these questionnaires provide a means to gain insight into the subjective component of affect in humans, they are an unparalleled tool in the translational study of affect.

Although describing positive animal welfare has been approached in numerous ways, for example, considering biological fitness (Barnett and Hemsworth, 1990; Hurnik and Lehman, 1988), physical condition (Broom, 1991; Veissier et al., 2008), or whether individuals are able to express natural behaviours (Hughes and Duncan, 1988; Veissier et al., 2008), ultimately, concerns about animal welfare stem from the belief that non-human animals subjectively experience negative affective states. Therefore, consideration of the psychological state of the animal should be of utmost importance when measuring welfare (Dawkins, 2008; Duncan, 2004).

In the United Kingdom, 1.89 million protected animals[2] were used for experimental procedures in 2016 (UK Home Office, 2017); these experimental procedures were considered to induce moderate suffering in 26% of animals used, and severe suffering in 5% of animals used (UK Home Office, 2017). Improving welfare is important for the benefit to the individual animals but compromised welfare is also considered to reduce the reliability and repeatability of scientific results (Poole, 1997; Prescott and Lidster, 2017). Therefore, ensuring positive laboratory animal welfare is of paramount importance. Accurate and reliable measures of animal affect would allow the impact of refinements to experimental procedures and husbandry in terms of animal welfare to be assessed. Moreover, improved measures of welfare could also help to progress a number of fields in which affect research is conducted, such as neuroscience, pain research, and psychopharmacology.

Physiology, behaviour and/or cognition are typically considered when measuring affect in non-human animals. Several physiological measures have been proposed to measure individual's emotional response to a stimulus or situation (Carter, 2001). These include measure of hormone levels (e.g. ACTH (Adrenocorticotropic hormone) or glucocorticoid) reflecting HPA (hypothalamic pituitary adrenal) axis activity (Abelson et al., 2007; Carter, 2001; Goldstein and Kopin, 2008; Moberg and Mench, 2000; Mormède et al., 2007), or heart rate variability or skin temperature reflecting autonomic nervous system activity (Moberg and Mench, 2000; Stewart et al., 2005;

[2]Where a protected animal is defined as 'any living vertebrate other than man and any living cephalopod'.

Von Borell et al., 2007), However, while heart rate and ACTH levels might rise in response to anxiety-inducing situations such as being pursued by a predator, they can also rise in positively valenced states of excitement (Ralph and Tilbrook, 2016). Hence, these physiological measures might reflect affective arousal as opposed to affective valence.

Methodological issues are also a major drawback of physiological measures of affect. For example, blood sampling is likely to be stressful for an individual and so, depending on the time course of the physiological response, the measure might largely reflect the emotional response to the sampling (Broom and Johnson, 1993). Caution must also be taken to account for diurnal and seasonal variation in hormonal measures (Mormède et al., 2007; Ralph and Tilbrook, 2016). Furthermore, there appears to be a dissociation between verbal report and some physiological indicators of affect in humans, which calls the reliability of physiological measure of affect into question (Lane et al., 1997; Stone and Nielson, 2001).

Potential physiological measure of longer-term affect (mood) include telomere attrition and hippocampal volume. These biomarkers purport to measure an individual's cumulative positive and negative experiences (Bateson and Poirier, 2019; Poirier et al., 2019). The use of such biomarkers is supported by research which has linked telomere attrition to anxiety and depression (Darrow et al., 2016), and a healthy lifestyle to reduced levels of telomere attrition (Puterman et al., 2015). Hippocampal atrophy has also been associated with putatively negative mood states in both human and (Arnone et al., 2016, 2013; McKinnon et al., 2009) and a range of non-human animals (Poirier et al., 2019; Sierakowiak et al., 2015). Moreover, rats that were exposed to stressors including bright light and loud sounds were found to have rostral expansion of sites in the nucleus accumbens involved in mediating avoidant or defensive behaviour, at the expense of sites that mediated approach or appetitive behaviour. The opposite was found to be true for rats housed in consistently familiar, dark, and quiet environments (Reynolds and Berridge, 2008).

Behavioural measures provide a less invasive measure of affect than physiological measures. Vocalisations and facial expressions have been proposed as measures of affective state in several species (Berridge, 1996; Darwin, 1872; Descovich et al., 2017; Finlayson et al., 2016; Knutson et al., 2002; Sandem et al., 2002). In rats, for example, 50kHz ultrasonic vocalisations (USVs) and a wider ear angle have been associated with positively valenced affect, while 22kHz USVs have been associated with negatively valenced affect (Finlayson et al., 2016; Knutson et al., 2002). However, vocalisations and facial expressions likely have a communicative function and so may be influenced by the presence or absence of conspecifics (Marler and Evans, 1996; Paul et al., 2005; Zimmerman et al., 2003) which would reduce their reliability as a of measure of affect.

Several behavioural tasks have been developed to measure longer-term affect. The forced swim test, and its derivative the tail suspension test, are among the most commonly used behavioural assays for investigating depression-like states in mice and rats (Cryan et al., 2002; Kato et al., 2007; Porsolt et al., 1977; Steru et al., 1985). In these

tasks, rodents (typically mice or rats) are faced with an inescapable threat of drowning or entrapment and the duration of immobility is measured. Immobility in the task was considered to measure hopelessness or despair, common features of depression, as the animal perceives escape attempts to be futile and because a broad range of antidepressants resulted in reduced immobility times (Cryan et al., 2002; Cryan and Mombereau, 2004; Kara et al., 2017). However, the interpretation of immobility as a proxy for helplessness has been questioned and immobility is now considered to reflect an adaptive and learnt response which allows the animal to conserve energy and hence prolong survival (Bogdanova et al., 2013; Borsini and Meli, 1988; Molendijk and de Kloet, 2015; Naudon and Jay, 2005; West, 1990). Although there is evidence to suggest that learning is impaired in depression, this new perspective on immobility weakens the validity of these tasks.

A less contentious alternative to measuring depression-like states is the sucrose preference test which measures hedonic capacity. However, the validity of this test has been challenged by the finding that humans with depression show no reduction in their preference for sucrose over water (Berlin et al., 1998; Dichter et al., 2010). Additionally, the extent to which non-human animals show a behavioural response, such as change in latency to approach a reward or consumption of a food reward following an increase or decrease in reward availability, deemed 'contrast effects', has also been proposed as a measure of welfare. More specifically, a greater behavioural change following a shift from smaller to a larger reward is associated with more a positively valenced affective state, while a greater behavioural change following a shift from a larger to a smaller reward is associated with a more negatively valenced affective state (Burman et al., 2008b; Neville et al., 2017).

The most frequently used behavioural assay to assess anxiety-like states in non-human animal models is the elevated plus maze (Montgomery, 1955; Pellow et al., 1985). This task measures the propensity of mice and rats specifically (as they are thigmotactic, Crozier and Pincus, 1928) to spend time in an open area relative to an enclosed area, and thereby pits exploration against safety. Individuals in anxiety-like states are expected to spend more time in the safe enclosed area. However, while administration of traditional anxiolytics can result in rats spending more time in closed areas, antidepressants that are used to treat generalised anxiety disorder in humans do not consistently produce changes in exploratory behaviour (Borsini et al., 2002; Carobrez and Bertoglio, 2005; Rodgers et al., 1997).

Other tasks developed to examine the relationship between affect and cognition in humans have also been adapted for non-human animals. Arguably, the two most extensively used affect-measuring cognitive tasks in humans are the visual dot probe tasks (MacLeod et al., 1986) and modified Stroop colour naming tasks (Williams et al., 1996). The former measures the extent to which presentation of threat words (e.g. 'pain') interferes with the detection of a dot probe, and the latter measures the extent to which naming the colour in which a word is written is influenced by the valence of the word itself (e.g. 'death' vs. 'chair'). Performance on these tasks has been shown

to be influenced by affect, with anxious individuals showing a greater decline in their ability to detect dots or name colours when threat words are shown in comparison to non-anxious individuals. While near-direct translations of these tasks exist for a range of non-human animals (Allritz et al., 2016; Cussen and Mench, 2014), at their core these tasks assess the extent to which affect biases attention, which has also been investigated in the context of animal welfare (Crump et al., 2018). In particular, the extent to which an individual attends to potential punishers appears to provide a measure of affect, although specifically high arousal (e.g. anxiety-like states) as opposed to low arousal (e.g. depression-like states; Bethell et al., 2012; Lee et al., 2018, 2016). In addition to biases in attention, biases in decision-making have also been used to investigate affect in both human and a range of non-human animals. This will be discussed in greater detail in section 1.3.

1.1.4 Affective neuroscience

Ultimately, the behavioural, cognitive, physiological, and subjective components of affect have a neural origin. In particular, the adrenergic, dopaminergic, GABA (gamma-Aminobutyric acid), and serotoninergic systems have been implicated as determinants of affect and also have a role in the processing of rewards and punishers (Kalia, 2005; Lövheim, 2012; Price and Drevets, 2010; Ruhé et al., 2007). Variation in the sensitivity of these systems resulting from either genetic or environmental factors may be responsible for variation in affect between individuals and underlie susceptibility to affective disorders in both human and non-human animals. Here, we will discuss each of these systems and their relationship to affect.

Epinephrine and norepinephrine are hormones and neurotransmitters that bind to adrenergic receptors. The adrenergic system is involved in the early stages of a stress response; it prepares the body for action by increasing heart rate, triggering the release of glucose, and increasing blood flow to skeletal muscles (Sapolsky et al., 2000). Brains of depressed human patients have reduced levels of norepinephrine, and antidepressant drugs such as reboxetine selectively-target the adrenergic system (Klimek et al., 1997; Massana, 1998). Genetic studies in both humans and mice have linked variation in norepinephrine transporter genes to susceptibility to depression-like states (Haenisch et al., 2009; Lin and Madras, 2006).

Dopamine is a neurotransmitter that can have both inhibitory and excitatory effects on target dopamine neurons. Dopamine is involved in moderating reward and punisher related behaviour (Berridge and Robinson, 1998; Ikemoto and Panksepp, 1999; Niv et al., 2007; Smith and Dickinson, 1998). In particular, dopamine activity has been shown to encode reward prediction errors (Schultz et al., 1997) and mediate the incentive salience (i.e. 'wanting') of rewards (Reynolds and Berridge, 2002). Dopaminergic-system dysregulation is associated with depression (Papakostas, 2006) and depletion of tyrosine (the precursor to dopamine) has been shown to alter reported mood in humans, although only in individuals at risk of major depressive disorder (Ruhé et al., 2007). Moreover, environmental enrichment, which is assumed

to induce a relatively positive affective state, can lead to changes in the mesocorticolimbic dopamine system in a range of non-human animals (Bezard et al., 2003; Laviola et al., 2008; Segovia et al., 2010; Zhu et al., 2005)

The neurotransmitter GABA is the major inhibitory neurotransmitter in the mammalian brain (Nemeroff, 2003). GABA is involved in eliciting appropriate responses to both appetitive and aversive stimuli (Reynolds and Berridge, 2002). Reduced GABA levels are associated with panic disorder (Goddard et al., 2001; Nemeroff, 2003). A number of commercially available treatments for anxiety disorders such as barbiturates, benzodiazepines, and gabapentins are thought to work by enhancing GABA function.

Serotonin is a primarily inhibitory neurotransmitter that binds to serotonergic receptors (Owens and Nemeroff, 1994). Serotonin is thought to guide behaviour in the face of aversive stimuli; and has been implicated in the prediction of aversive events, as well as freezing and withdrawal when presented with punishers (Boureau and Dayan, 2011; Owens and Nemeroff, 1994). There is a wealth of evidence indicating a link between low levels of serotonin and depression (see Owens and Nemeroff, 1994). Antidepressant drugs that specifically target the serotonergic system, such as citalopram and fluoxetine, are commonly prescribed to humans experiencing depression (Mars et al., 2017). Depletion of serotonin by dietary depletion of its precursor tryptophan in human subjects can reduce mood in those predisposed to depression (Ruhé et al., 2007) and post-mortem examination of brain tissue from depressed and suicidal human patients revealed reduced concentrations of serotonin (Kambeitz and Howes, 2015). Knockout mice lacking the serotonin 1a receptor exhibit enhanced anxiety-related behaviour (Ramboz et al., 1998).

Importantly, these neurotransmitter receptor systems are highly conserved across species (Jorgensen, 2014; Ottaviani and Franceschi, 1996; Venter et al., 1988). Consequently, investigating whether pharmacologically altering levels of affect-modulating hormones or neurotransmitters in non-human animals induces the predicted changes in a measure of affective state, given its effect on human mood, has become the 'gold' standard for validating measures of affect in non-human animals. In Chapter 2, we describe and discuss the results of a meta-analysis and systematic review of pharmacological manipulations on judgement bias, a potential measure of affect (see section 2.3).

1.2 Affect, cognition, and decision-making

1.2.1 Affect and cognition

Cognition has been defined as 'the ways in which animals take in information through the senses, process, retain and decide to act on it' (Shettleworth, 2000). Beck's Schema theory suggests that distortions in cognition contribute to the development and maintenance of affective disorders (Beck and Alford, 1967; Beck, 2002). These distortions,

such as overgeneralising, catastrophising, and magnification of the importance of negative outcomes, lead to negative views about the self, the world, and the future (Beck and Alford, 1967; Beck, 2002). Similarly, it has been suggested that an individual's appraisal of a situation or event determines the resulting affective state; for example, a snake might only induce fear if an individual considers the snake to pose a threat (Lazarus, 1991).

Affect has also been proposed to influence various aspects of cognition including attention, memory, and decision-making. The influence of emotion on attention is clear; potentially fear-inducing stimuli, such as snakes or spiders, rapidly grab our attention (Eastwood et al., 2001; Öhman et al., 2001). Anderson (2005) also demonstrated that humans have an enhanced ability to detect emotionally salient words relative to neutral words. Furthermore, threatening words distract anxious humans and sad words distract depressed humans to a greater extent than clinically healthy humans. (Williams et al., 1996). Similarly, in humans, anxiety is associated with greater attention to threatening words in the dot probe task (MacLeod et al., 1986), while positive mood is associated with greater attention to reward-related words (Tamir and Robinson, 2007). However, these results are complicated by the finding that humans will gaze at happy faces for a longer duration following a negative mood induction, which has been proposed as a form of mood repair (Sanchez et al., 2014).

Affect has also been proposed to bias memory recall. For example, mood-congruent memory describes the finding that memories are more readily retrieved when their emotional significance is consistent with the individual's current affective state (Blaney, 1986; Elliott et al., 2002). Likewise, affect is considered to influence interpretation of ambiguous stimuli; when presented with homophones (e.g. die/dye) anxious humans have a greater tendency to interpret the word negatively than non-anxious individuals (Eysenck et al., 1991; Mathews et al., 1989) and when presented with neutral faces, humans experiencing more negative mood states are more likely to interpret the face as being sad (Hale III et al., 1998; Leppänen et al., 2004).

Several theories have sought to explain the influence of affect in decision-making, such as the affect heuristic (Finucane et al., 2000) and somatic-marker hypothesis (Bechara and Damasio, 2005; Damasio, 1994), both of which describe how affective responses may guide decision-making. Most commonly, negative affective states have been associated with more risk-averse decision-making. For instance, depressed humans score higher on self-reported measure of harm avoidance (Joffe et al., 1993) and following a negative affect induction humans report lower tendencies for risk-taking (Yuen and Lee, 2003). Additionally, Anderson et al. (2012) found that anxious humans were more likely to make decisions that allowed punisher avoidance as opposed to reward gain. Paul et al. (2011) found a positive correlation between positive affect scores (as measured by the PANAS questionnaire) and bias towards risk-seeking decision-making, and likewise a negative correlation between negative affect scores and bias towards risk-seeking decision-making in humans. However, negative affect

has also been associated with increased risk-taking behaviour, particularly with regards to gambling (Koot et al., 2013; Porcelli and Delgado, 2009; Preston et al., 2007).

1.2.2 How are decisions made?

Historically, little consideration has been given to the influence of affect on decision-making processes. However, the question of how an optimal decision-maker should behave, and how individuals deviate from this, especially in the context of monetary rewards and losses, has long been studied. Expected value, first considered in the 17^{th} century (Daston, 1980; Ore, 1960), quantifies the long-run average value of a choice as the summed probability of each possible outcome resulting from that choice multiplied by its value, hence providing a means for options to be compared and decisions to be made. However, as Bernoulli (1738) recognised, expected value does not allow for individual differences in the subjective value of the outcome (Bernoulli, 1954; Stearns, 2000). Expected utility theory considers that the subjective value of an outcome may differ from the absolute value, providing an explanation as to why individuals do not necessarily select options with the highest expected value (Bernoulli, 1954; Stearns, 2000; Von Neumann and Morgenstern, 1953). Prospect theory, derived from economic experiments and observations of human behaviour, further accounts for deviations in behaviour from the predictions of expected value theory. It considers that individuals have a biased weighting of probabilities, such that the subjective probability of high probability events is reduced and subjective probability of low probability events in increased relative to the true probability. Prospect theory also allows for asymmetry in the value of outcomes, such that losses are weighted more heavily than gains.

A commonality between these theories is that they posit that individuals integrate information about the probability and value of the potential outcomes of a specific decision, thereby yielding a single value which provides a common currency for decision-making. Neurobiological evidence suggests that such a value is estimated and used to inform decision-making (Breiter et al., 2001; Knutson et al., 2005; Lau and Glimcher, 2008; Padoa-Schioppa and Assad, 2006, 2008; Platt and Glimcher, 1999). It has been demonstrated that the expected value of a choice modulates the activity of intraparietal neurons in rhesus macaques (*Macaca mulatta*), and that this activation correlates with option selected by the macaque (Platt and Glimcher, 1999). Additionally, human fMRI studies have revealed activation in mesolimbic brain regions that correlate with the expected value of a choice (Breiter et al., 2001).

In addition to an integrated representation of reward value and probability, these two components of expected value have independent neural representations (Knutson et al., 2005). The mesolimbic dopaminergic system has been implicated in the valuation of reward-related stimuli, specifically determination of the incentive value of rewarding stimuli (Berridge, 2012; Berridge and Aldridge, 2008; McClure et al., 2003; Zhang et al., 2009). The incentive salience of a reward (the extent to which it is desired) has been shown to depend on D2 receptor activity (Maldonado et al., 1997; Trifilieff et al., 2013). It has been proposed that the hedonic value of a reward

is encoded by the opioid system (Berridge et al., 2009; Peciña and Berridge, 2000). Encoding of probability estimation appears to also involve modulation of dopaminergic activity, with the phasic activation of dopamine neurons exhibiting variation dependent on the probability with which a reward occurs (Fiorillo et al., 2003; Tobler et al., 2007).

1.2.3 Affect and probability estimates

Numerous studies have demonstrated that affect can alter subjective probability assessments in humans. In particular, positive affect appears to result in greater estimates of the probability of rewarding events and lower estimates of the probability of punishing events, while negative affect appears to have an opposite effect on probability estimation (Johnson and Tversky, 1983; Nygren et al., 1996; Wright and Bower, 1992). For example, reading a newspaper story about death, regardless of the type of death reported (e.g. homicide, illness, or accident) both induced negative affect and increased the individual's estimate of the frequency of fatal accidents (Johnson and Tversky, 1983). Conversely, a positive affect induction was shown to increase an individual's rating of their likelihood of winning a bet (Nygren et al., 1996).

Affective disorders are also associated with biased probability estimates. Depressed and anxious humans rate the probability that they will experience punishing events more highly than clinically healthy individuals (Andersen et al., 1992; Butler and Mathews, 1983; MacLeod and Byrne, 1996; Muris and Van der Heiden, 2006. However, there is some evidence that greater estimation of punishing events is more closely associated with anxiety than depression (Muris and Van der Heiden, 2006) and that depression is instead linked to reduced estimation of the probability of rewarding events (MacLeod and Byrne, 1996; Muris and Van der Heiden, 2006). Furthermore, when given the task of generating scenarios of potential future events, humans with anxiety generate a greater number of negative event scenarios in comparison to clinically healthy humans (MacLeod and Byrne, 1996), while depressed humans generate fewer positive events (Bjärehed et al., 2010; MacLeod and Byrne, 1996; MacLeod and Salaminiou, 2001). An experiment was conducted in which individuals with varying degrees of depression were asked to report both the probability of an event occurring and at a later date were asked whether the event occurred (Strunk et al., 2006). It was found that individuals with more severe depressive symptoms had more pessimistic probability estimates and that their predictions about the future were imprecise. However, as discussed in section 1.1.2, a greater number of 'false' alarms may be adaptive for depressed and anxious individuals, where such states arise from previous experience of rewards and punishers.

1.2.4 Affect and reward and punisher valuation

There is contradictory evidence regarding the influence that affective valence might exert on reward and punisher valuation. In a range of non-human animals, manipulations that are anticipated to induce negatively valenced affect such as chronic stress or early-life adversity have been associated with both an increased and a decreased valuation of rewards (Kleen et al., 2006; Neville et al., 2017; Rüedi-Bettschen et al., 2006; Shaham and Stewart, 1994). Environmental enrichment, assumed to induce a relatively positive affective state, has been shown to reduce self-administration of amphetamines (Bardo et al., 2001; Green et al., 2002) and reduce the hedonic value of a number of drugs including cocaine, nicotine, and heroin (Green et al., 2003; Solinas et al., 2008). Similarly, in humans, although anhedonia is a core symptom of depression, hyperphagia can also be observed (American Psychiatric Association, 2013).

Research has shown that depression is associated with a preference for low-magnitude rewards obtained with low-effort compared to high-magnitude rewards obtained with high effort (Treadway et al., 2012, 2009). Thus, it has also been proposed that affective disorders, specifically depression, modulate the value of the effort required to obtain a reward, in addition to value of the reward itself. This has been further supported by the finding that escitalopram, an antidepressant, modulates effort expenditure, as measured by force exerted on a handgrip (Meyniel et al., 2016).

With regards to punisher valuation, it has been demonstrated that anxiety-like states increase the physiological response to potential punishers (Davis et al., 2010; Grillon, 2008) which could reflect that anxiety increases the subjective value of the punisher. However, despite anxiety and depression often being comorbid in humans, a meta-analysis identified that major depressive disorder is associated with a reduced reactivity to punishers (Bylsma et al., 2008).

The temporal dynamics of the affective state might also modulate reward and punisher valuation; there is evidence for an increased reward valuation in shorter-term negatively valenced affective states, but a decreased reward valuation in longer-term negatively valenced affective states (Spruijt et al., 2001).

This uncertainty about how affect should influence reward and punisher valuation also translates to hypotheses regarding the putative function of the relationship between affect, or reward and punisher experience, and subjective value. Reward utility functions are considered to be concave (i.e. increasingly less steep), suggesting that the additive subjective value of a reward should decrease with increasing gains; an individual that is starving should value food more highly that an individual that is sated (Hsee and Rottenstreich, 2004; Hsee et al., 2005). Yet, in reward-sparse environments a high reward value which could promote greater activity would be antithetic to the conservation of energy in the face of limited resources. Trimmer et al. (2017) proposed that punisher experience should determine the value of rewards, specifically food rewards; food should be valued more highly in environments in which threats are more frequent as maintaining the high energy reserves needed to evade predation

becomes more important. Likewise, it has been suggested that an increased valuation of rewards in more negative mood states could function to regulate mood by motivating the individual to seek mood-enhancing rewards (Morris and Reilly, 1987; Sanchez et al., 2014).

1.3 Judgement bias: decision-making as a measure of welfare?

The judgement bias task, first described in 2004 by Harding et al. and based on discrimination learning, has been used to measure the cognitive component of affect, specifically altered decision-making, in numerous species, including rats (*Rattus norvegicus*; Brydges et al., 2011; Burman et al., 2008a; Harding et al., 2004), starlings (*Sturnus vulgaris*; Brilot et al., 2010; Matheson et al., 2008), humans (Iigaya et al., 2016; Paul et al., 2011), and bees (*Apis mellifera*; Perry et al., 2016). In a common variant of this task, an individual is trained to make different actions in response to sensory stimuli that differ detectably on a sensory continuum (e.g. high and low frequency tones). There is a safe action which leads to no reward and no punishment, and a risky action which leads to either a reward or a punisher[3], the outcome of which is determined by whether the predictive sensory stimulus is at one or the other end of the trained sensory continuum. During testing, the individual is required to make one of these two actions in response to intermediate ambiguous sensory stimuli (Fig. **??**). More risk-seeking (often deemed 'optimistic') individuals, those who execute the risky action or do so with greater alacrity, are considered to have a greater expectation of rewards, or reduced expectation or punishers, and thus are considered to be experiencing a more positively valenced affective state in comparison with individuals that are more risk-averse (often deemed 'pessimistic'), those who execute the safe action or execute the risky action more slowly (Bateson et al., 2011; Mendl et al., 2010; Nettle and Bateson, 2012; see Fig. 1.3). The hypothesis that relatively risk-seeking behaviour on the judgement bias task is associated with a relatively positive affective state has overall been supported by studies which have involved pharmacological and environmental manipulations of affect, such as the provision of enrichment and administration of antidepressant drugs, although opposite and null results have been observed. In Chapter 2, we more formally examine this hypothesis by conducting a meta-analysis and systematic review of pharmacological manipulations of judgement bias in non-human animals.

[3]Here, following Kacelnik and Bateson (1996) and Niv et al. (2012), we are defining 'risk' as outcome variability such that actions that have greater variability in their potential outcome are considered to be more risky

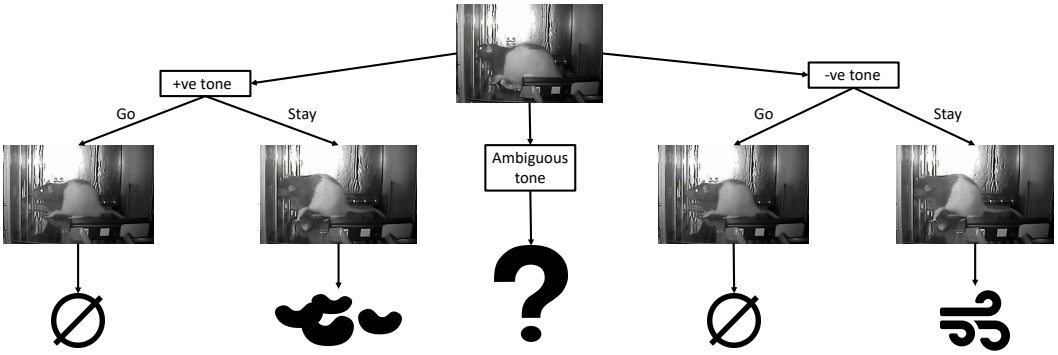


FIGURE 1.2: An example of judgement bias test, using the task described by Jones et al. (2018): following trial initiation, the rat is presented with a tone which is either +ve, -ve, or ambiguous. If they execute the risky response ('stay') when the +ve tone is presented they will be rewarded with food, but executing this response when the -ve tone is presented will result in a punishing air-puff. The safe response ('go') results in no reward or punisher. During testing, the rat is presented with an ambiguous tone and must decide whether to execute the 'stay' or 'go' response

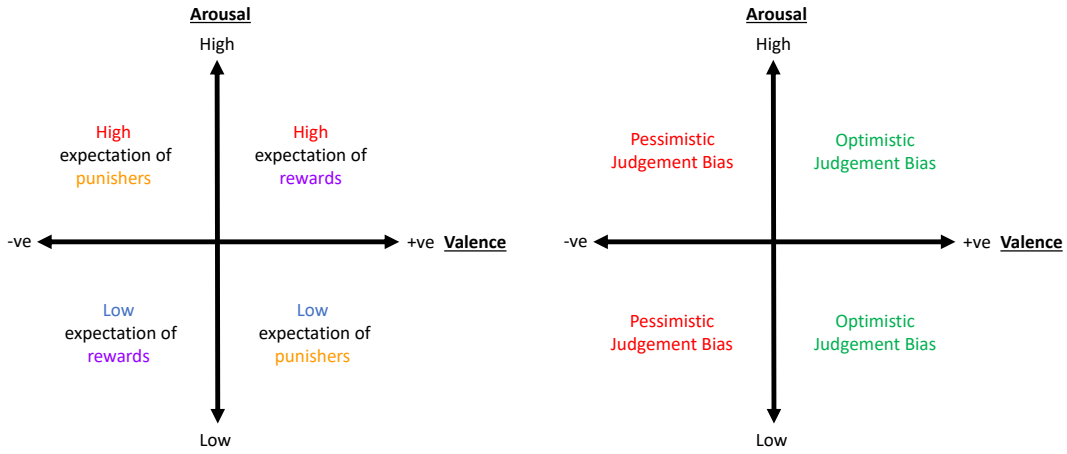


FIGURE 1.3: The theoretical framework underlying the judgement bias task: (A) positive affective valence is associated with either a high expectation of rewards or low expectation of punishers (depending on affective arousal), and likewise negative affective valence is associated with either a low expectation of rewards or high expectation of punishers (depending on affective arousal); (B) this in turn leads to affect-dependent predictions about an individual's judgement bias relative to other individuals completing the same task.

The judgement bias task has several strengths as a measure of affect; interpretation of judgement bias results does not require species-specific knowledge, it purports to measure affective valence as opposed to affective arousal, and human-versions of the task have found that judgement bias correlates with subjective reports of affect (Anderson et al., 2012; Iigaya et al., 2016; Paul et al., 2011).

However, one major drawback of the task is the typically lengthy period of training of non-human animal subjects required before testing can commence. While this problem was recently addressed by Jones et al. (2018) who developed an automated task which reduced the duration of training in rats, in Chapter 7 we propose and examine the validity of a novel measure of affect with a reduced training period. Limitations also exist in the interpretation of judgement bias results. Firstly, judgement bias only provides a relative and not absolute measure of affect; an individual that is overall risk-seeking is not necessarily in a positively valenced affective state. This is because methodological differences, such as the rewards or punishers used and reinforcement schedule, can also influence judgement bias (Jones et al., 2018). Secondly, there has been a paucity of research examining the cognitive mechanisms underlying the relationship between affect and judgement bias. As outlined in section 1.2, affect could influence decision-making via an altered probability estimation and/or an altered subjective valuation of the outcomes, with the latter often receiving little consideration in the context of judgement bias. This complicates interpretation of judgement bias data as the absence of strong evidence on which to base predictions about how an affect manipulation will alter reward or punisher valuation presents a challenge when making conclusions about how affect relates to judgement bias. To address this, in Chapters 3, 4, 5, and 6 we use computational modelling to disentangle the influence of an altered probability estimation and an altered subjective valuation on decision-making in the judgement bias task.

Biases might also arise through other cognitive processes such as attention or perception (Mendl et al., 2009). For example, given that negative affect is associated with greater attention to threatening stimuli (MacLeod et al., 1986; Williams et al., 1996), and positive affect with greater attention to positive stimuli (Tamir and Robinson, 2007), responses may be more rapid when the valence of the stimulus presented is more congruent with the valence of the individual's affective state (Mendl et al., 2009). Attention-related biases in judgement may be most apparent where response latency is the outcome measure (Mendl et al., 2009). As the focus of this thesis is specifically on the decision-making process, we will solely examine choice and not response latency. Affect might also alter an individual's perception of the stimuli; there is some evidence that affective valence can influence perception, for example, anxious humans are more likely to perceive neutral facial expressions as negative (Blanchette et al., 2007). Consequently, affect may cause the ambiguous stimuli to be perceived as indicative of the negative or positive outcome, which may be most evident at the ambiguous stimuli closest to the reference cues (Mendl et al., 2009). However, the possibility that affect exerts a top-down influence on perception has been challenged (see Firestone and Scholl, 2016).

1.4 Computational modelling

1.4.1 What is computational modelling?

Marr (1982) proposed three-levels at which the brain can be understood: the computational, the algorithmic, and the implementational. In terms of cognition, at the computational level the precise problem is specified, at the algorithmic level the computational problem is solved, and at the implementational level the neurobiological substrate involved in solving the algorithm is considered. Computational modelling of behavioural data may involve all three levels.

The aim of computational modelling is to recreate data such as reaction time, choice, or neural firing, using a formal and mathematical description of the processes through which the data arose. A key feature of computational modelling is that it is generative and not just descriptive; a computational model generates its own data to be compared with experimental data. Computational models can be used for identifying the parameter(s) which best account(s) for the data and/or for identifying the model which best accounts for the data (Daw et al., 2011a; Wilson and Collins, 2019). Consequently, computational modelling can help to elucidate the cognitive processes underlying behavioural data by providing a clear and specific description of these processes (Daw et al., 2011a). In this thesis, we will use computational modelling as a tool to better understand the processes underlying decision-making.

1.4.2 Reinforcement learning

Reinforcement learning characterises how an individual learns through interactions with the environment and how they can make decisions which maximise reward intake while minimising punisher exposure (Bellman, 1952; Sutton and Barto, 1998). Markov decision processes (MDP) allow a formal description of problems to be solved through reinforcement learning (Bellman, 1957; Sutton and Barto, 1998). In particular, a MDP is defined as comprising states, actions, transition probabilities, and a reward function.

To illustrate this, consider a rat in a maze searching for a sucrose pellet. We can formalise this as a MDP; the states can be defined as all possible locations in the maze, and the actions can be defined as the movements that rat make to traverse through the maze (e.g. left, right, forward, backward) to transition to new states. The transitions between the states could be deterministic where the action that the rat makes dictates the rat's new state (e.g. they execute action 'left' and subsequently move left), but if the maze was unsteady, the transitions could be probabilistic where the action and probability that the rat is thrown off course in any direction determine the new state (e.g. they execute action 'left' and move left with a probability of 0.7, and move right, forward, or backward each with a probability of 0.1). In this example, the reward is zero in all states except for the location of the sucrose pellet. However, we have assumed that the rat knows which state they occupy. A partially observable Markov decision process (POMDP) is a generalisation of a MDP in which the true

state is unknown but observations (e.g. of odours within a maze) allow a probability distribution over the possible states to be maintained (Bellman, 1957; Sutton and Barto, 1998). In Chapters 3, 4, and 6, we define decision-making in the judgement bias task as a POMDP.

Reinforcement learning can be either model-free or model-based (Dayan and Berridge, 2014; Doll et al., 2012; Gläscher et al., 2010). In model-free reinforcement learning, an individual takes a trial-and-error approach to learn how to act optimally, in which reward values are updated through experience and direct knowledge of the action-outcome contingency is not required. Decision-making depends on the actions which an individual has learnt results in the greatest expected value. Using the example of a maze, a rat using a model-free strategy may navigate the maze using the learnt value of each action, for example, turning left because experience has taught the rat that turning left has the greatest value. Hence, model-free learning does not make use of the transition probabilities and is retrospective. Prominent examples of model-free learning include the Rescorla-Wagner model (Rescorla et al., 1972; Sutton and Barto, 1998) and temporal-difference (TD) learning which encompasses Q-learning (Sutton, 1988; Sutton and Barto, 1998). In these models, learning is driven by a reward prediction error; the difference between an expected value estimate and the outcome (Rescorla-Wagner model) or updated expected value estimate at each time-step (TD learning). Hence, reward prediction error is central to both these (and other) model-free methods (Dayan and Berridge, 2014; Dayan and Niv, 2008; Gläscher et al., 2010; Sutton and Barto, 1998).

In model-based reinforcement learning, an individual maintains a model of the environment which they use to evaluate all possible outcomes and guide behaviour (Dayan and Berridge, 2014; Doll et al., 2012; Gläscher et al., 2010). A rat using a model-based strategy may navigate a maze according to the long-term value of future actions and transitions; considering the value of each action (e.g. 'left') according to where the action could ultimately lead them. Model-based learning involves prospective planning and is informed by transition probabilities. Value-iteration and policy-iteration are both model-based algorithms which can be used to solve MDPs (Bellman, 1952; Sutton and Barto, 1998).

Model-free and model-based approaches differ in their flexibility and computational demand; while model-free learning is less flexible than model-based learning but computationally non-intensive, model-based learning is fast and flexible but computationally intensive (Dayan and Berridge, 2014; Doll et al., 2012). There is evidence for both model-free and model-based learning in both humans and rats (Daw et al., 2011b; Doll et al., 2012; Gläscher et al., 2010; Hasz and Redish, 2018). In particular, goal-directed behaviour which is planned and purposeful such as pressing a lever with the goal of obtaining a sucrose pellet is often attributed to model-based learning, while stimulus-response habits such as pressing a lever even when the sucrose has been devalued ('going into autopilot') are attributed to model-free mechanisms

(Dayan and Berridge, 2014; Doll et al., 2012; Gläscher et al., 2010). Dopaminergic activity has been proposed to encode reward prediction errors; while there is no change in dopamine neuron firing following predicted rewards, unpredicted rewards lead to increased firing, and predicted but absent rewards lead to decreased firing (Glimcher, 2011; Montague et al., 1996; Schultz et al., 1997). Dopamine has also been implicated in model-based learning (Sharp et al., 2015). Experimental alterations of dopamine in humans have demonstrated that reduced levels of dopamine are associated with a greater tendency to use model-free as opposed to model-based learning (De Wit et al., 2012; Wunderlich et al., 2012).

1.4.3 Bayesian inference

Bayesian inference has also been applied to learning problems (Adams et al., 2016; Friston, 2008; Knill and Pouget, 2004; Parr et al., 2018). Bayes' rule states that the probability of an event occurring in the face of new evidence (posterior probability) depends on both a prior probability, which is the probability of the event prior to new evidence, and likelihood of the event, which is the probability of the new evidence given that the event will occur. Consequently, Bayes' rule can be used to determine the probable outcome of an action or likelihood of a belief about the state of the environment given sensory evidence from the environment. The mean and precision provide sufficient statistics for encoding the prior and likelihood probability distributions, and hence approximating Bayesian inference (Adams et al., 2016; Friston, 2008; Parr et al., 2018). More specifically, the posterior probability will lie between the means of the prior and likelihood distributions, with the proximity to either distribution governed by the precision of the distributions (assuming a Gaussian distribution for the prior and likelihood). The precision hence determines the extent to which new evidence influences current beliefs.

The Bayesian brain hypothesis states that our nervous system constructs a probabilistic model of the world that continuously updates as sensory information is received in a Bayesian manner (Dayan et al., 1995; Knill and Pouget, 2004). Moreover, the theory of predictive coding (Friston, 2008; Rao and Ballard, 1999), for which there is some biological evidence (Bastos et al., 2012; Shipp et al., 2013), suggests that perception is influenced by top-down predictions from a model of the environment to aid the efficient filtering of irrelevant information.

1.5 Computational psychiatry

1.5.1 What is computational psychiatry?

Computational psychiatry uses computational techniques to characterise and understand affective disorders. The essential premise of computational psychiatry is that affective disorders arise from or induce a deviation from optimal behaviour, and that by computationally defining optimality we can examine the specific suboptimalities

associated with affective disorders (Huys et al., 2016; Montague et al., 2012; Stephan and Mathys, 2014). It has been proposed that suboptimal decision-making can originate from: abnormalities in how the problem to be solved is defined, making wrong inferences (e.g. about state or actions), and by solving the right problem correctly but assuming an aversive environment (Huys et al., 2015). It must be noted that it is unclear whether seemingly suboptimal decision-making in individuals with affective disorders is truly suboptimal as such behaviour may serve an adaptive function, as outlined in section 1.1.2.

1.5.2 A computational perspective on the function of affect

Affective valence has been suggested to function similar to a Bayesian prior, informing an individual's estimation of the probability that their actions will be rewarded or punished to guide decision-making (Mendl et al., 2010; Trimmer et al., 2013). This proxy Bayesian prior would update following reward and punisher experience and hence reflect environmental conditions (Mendl et al., 2010; Mendl and Paul, 2017; Trimmer et al., 2013). This would obviate tracking of multiple historical rewarding and punishing experiences (i.e. encoding and recalling each individual rewarding and punishing experience) and allow more efficient encoding of past experience (Eldar et al., 2016).

Predictability is thought to be a central facet of affect (Clark et al., 2018; Huys and Dayan, 2009). Rewards can be more efficiently exploited, and punishers more efficiently avoided when they are predictable, and therefore more predictable environments should be associated with more positive affect (Huys and Dayan, 2009). Moreover, it has also been proposed that longer-term affect (mood) reflects a Bayesian hyperprior over uncertainty about outcomes; a belief about the extent to which the environment is predictable (Clark et al., 2018). In particular, the precision of this hyperdistribution is thought to underlie affective disorders. A greater precision around the belief that the environment is unpredictable would lead to a more persistent belief and rejection of any evidence that the world is stable, as observed in depressed individuals (Clark et al., 2018; Huys and Dayan, 2009). Indeed, humans with major depressive disorder that have higher Beck Helplessness Scores (BHS) perceive themselves to have a low level of control, as determined by a broader prior distribution about the outcome of their actions (Huys et al., 2009). In contrast a weaker precision, associated with anxiety, leads to greater action to resolve uncertainty about the predictability of the environment (Clark et al., 2018; Cornwell et al., 2017; Huys and Dayan, 2009).

However, this theoretical framework for affect does not account for recent evidence that reward prediction errors, and hence greater unpredictability, induce positive affect. It has been demonstrated that humans report greater happiness when they receive a reward that is greater than expected (Rutledge et al., 2014). Moreover, weather and sports results that are greater than expected are associated with greater

gambling suggesting a link between positive prediction errors and 'optimistic' decision-making (Otto et al., 2016). Considering this evidence, Eldar et al. (2016) suggested that affect reflects the cumulative impact of positive and negative prediction errors, purported to allow quick adaptation to changing environmental conditions. According to Eldar et al. (2016), affective disorders are thus considered to arise from suboptimalities in the expectation of reward and punishers; failure to update expectations in worsening conditions would lead to the continual generation of negative prediction errors and hence induce a depressed mood.

Interestingly, depression has been associated with altered prediction error signalling, specifically attenuated signals in hippocampus, rostral anterior cingulate cortex, striatum, and thalamus, and stronger signals in the amygdala (Gradin et al., 2011; Kumar et al., 2008). Moreover, the magnitude of this attenuation and amplification correlated with the severity of the depression (Kumar et al., 2008). Stress induction in healthy subjects has also been observed to modulate the strength of the prediction error signal (Robinson et al., 2013).

1.5.3 How does affect modulate the processes underlying decision-making?

Computational approaches have proved highly effective and fruitful in examining the cognitive processes underlying affect-induced changes in decision-making behaviour. Studies have investigated the relationship between affect and probability estimation. For example, it was demonstrated that more optimistic individuals, as determined by the revised life orientation test, had a greater prior belief that they would be rewarded in a probabilistic task (Stankevicius et al., 2014). Another study attributed suboptimal behaviour related to stress to a reduced estimation of the quality of the environment (Lenow et al., 2017).

Computational modelling has also been used to better understand the relationship between affect and reward and punisher valuation. Using a reinforcement learning model, a meta-analysis of human behavioural data from a probabilistic reward task revealed that major depressive disorder was associated with a reduced reward sensitivity (Huys et al., 2013). It was also found that apathy in human depression is associated with an increased effort cost as opposed to a reduced reward valuation (Bonnelle et al., 2015). In relation to anxiety and subjective valuation, Charpentier et al. (2017) found that humans with an anxiety disorder did not differ from clinical healthy individuals in their relative valuation of losses to gains but were more averse to uncertainty about the outcome of their actions. Analysis of human judgement bias data using a Bayesian decision-theoretic model found that a positive affect induction both increased the participants' value of a potential reward relative to the potential loss and biased them towards the risky choice (Iigaya et al., 2016).

There is also evidence from computational analyses to suggest that affect alters the extent to which individuals employ model-free and model-based reinforcement learning to solve problems. For example, humans experiencing high levels of chronic

stress had a reduced tendency for model-based learning when confronted with an acute stressor (Radenbach et al., 2015). Similarly, model-free learning models were found to better explain the behaviour of depressed humans than non-depressed humans on a task which pitted exploration against exploitation (Blanco et al., 2013). Additionally, Huys et al. (2012) found that humans that experienced greater levels of depression (as determined by Beck's depression inventory) had a greater tendency to dismiss further evaluation of a potential strategy when presented with a large loss, even though this was counterproductive, in a task which required them to construct a model-based representation of the environment (Huys et al., 2012).

Affect has also been proposed to influence learning. The extent to which an individual updates their belief about the environment when given new information should depend on environmental volatility; faster learning should be required in unstable environments compared with those with little fluctuation (Browning et al., 2015). However, anxiety in humans has been associated with a reduced ability to modulate the speed of learning in accordance with environmental volatility (Browning et al., 2015). Although a number of studies have suggested that depression is associated with reduced learning in humans (Chase et al., 2010; Dombrovski et al., 2013), a number of studies have found no evidence for this (Dombrovski et al., 2010; Gradin et al., 2011; Huys et al., 2013), and so the effects of affective state on learning are unclear.

1.5.4 Computational psychiatry in non-human animals

Despite the importance of investigating affect in non-human animals, both for improving animal welfare and to better understand human affective disorders, there remains a paucity of research that uses computational analyses to investigate affect-induced behavioural changes in non-human animals. One example of where computational modelling has been instructive in examining affect in non-human animals a study which applied drift-diffusion modelling to rat judgement bias reaction time data (Hales et al., 2016). The drift-diffusion model considers that evidence about the stimulus presented is accumulated until a decision threshold is reached, with the key parameters being the speed at which evidence is accumulated, the starting point of the evidence accumulation, and distance from the threshold. It was found that pharmacologically induced negative affect in rats increased the threshold for making the risky-response (Hales et al., 2016). Additionally, consistent with the human literature, chronic stress in rats has been associated with a greater likelihood of employing model-free learning (Dias-Ferreira et al., 2009).

1.6 Aims of the thesis

As outlined in this review, there is still much uncertainty about the adaptive significance of affective states and the relationship between reward and punisher experience, affect, and decision-making. The overarching aims of this thesis are to examine which

(if any) specific components of reward and punisher experience inform decision-making (e.g. their frequency and predictability), the cognitive processes underlying judgement bias (e.g. reward and punisher valuation and expectation) and their relationship with reward and punisher experience, and the extent to which the relationship between reward and punisher experience and decision-making is consistent between rats (a commonly used laboratory animal with whom the original judgement bias study was conducted) and humans (with whom we can also investigate how reward and punisher experience relates to subjectively experienced affect). It is hoped that this will inform our understanding of judgement bias as a measure of welfare and more generally contribute to the fields of animal welfare science and computational psychiatry.

We started by conducting a systematic review and meta-analysis of pharmacological manipulations of affect to assess the validity of the judgement bias task as a measure of affect in non-human animals (Chapter 2). Then, we developed human versions of an automated rat judgement task (Jones et al., 2018) and examined how human decision-making and self-reported affect are influenced by primary (i.e. food) and secondary (i.e. monetary) reward and punisher experience (Chapters 3 and 4). Next, we conducted rat judgement bias studies in which we manipulated reward and punisher experience prior to and within testing (Chapters 5 and 6). Finally, we assessed whether decision-making in a task which pits exploration against exploration could provide a novel measure of affect in rats (Chapter 7).

Computational psychiatry has arguably been instrumental in furthering our understanding of human affective disorders. Hence, the experimental chapters of this thesis describe a computational approach to data analysis to aid in the elucidation of the cognitive processes underlying behaviour. In Chapters 3, 4 and 6 we report a novel computational model of decision-making on the judgement bias task which we developed.

Chapter 2

Pharmacological manipulations of judgement bias: a systematic review and meta-analysis

Chapter summary: Validated measures of animal affect are crucial to research spanning a number of disciplines. Judgement bias, which assesses decision-making under ambiguity, is a promising measure of animal affect. One way of validating this measure is to induce affective states using pharmacological manipulations and determine whether the predicted judgement biases are observed. Here, we conducted a systematic review and meta-analysis using data from 20 published research articles that use this approach, from which 557 effect sizes were extracted. The results of the meta-analysis suggest that pharmacological manipulations overall altered judgement bias at the probe cues as predicted. However, there were several moderating factors including the neurobiological target of the drug, whether the drug was expected to induce a relatively positive or negative affective state, dosage, and the presented cue. This may partially reflect interference from adverse effects of the drug, such as sedation. Thus, while judgement bias can be used to measure pharmacologically induced affective states, potential adverse effects of the drug should be considered when interpreting results.

2.1 Introduction

Measurement of affective state, which is defined as comprising both short-term emotions and longer-term moods (Mendl et al., 2010), is important to a number of disciplines including psychopharmacology, neuroscience, and animal welfare science, as well as being of societal interest. For example, mood disorders are a significant global concern; it is estimated that 780,000 people died by suicide in 2015, with on average one death every 40 seconds (World Health Organization, 2017). Major depressive

disorder is ranked as the largest single contributor to global disability, and anxiety disorders are ranked sixth (World Health Organization, 2017). The development of pharmacological treatments for mood disorders has been largely dependent on empirical studies using non-human animals (Rupniak, 2003; Valvassori et al., 2013). Reliable and validated measures of affective state in non-human animals are therefore crucial to understanding the neurobiological aetiology of these disorders and to assist in the development of novel treatments. In particular, measures should have both predictive validity (i.e. the extent to which the measure is altered in the predicted direction by drugs which alter human affect) and construct validity (i.e. the extent to which they measure precisely what they claim to measure) (Nestler and Hyman, 2010). Predictive validity is typically regarded as the 'gold-standard' for validating novel behavioural measures of affective state (De Vry and Schreiber, 1997; McArthur and Borsini, 2006).

The outlined deficiencies in currently used assays (e.g. forced swim test, sucrose consumption, elevated plus maze, see Chapter 1) means that there is a clear need for improved methods to measure affective state in non-human animals that have both construct and predictive validity. The judgement bias task (sometimes referred to as the cognitive bias task or ambiguous cue interpretation task) provides an alternative means to examine affect in non-human animals and has been used widely in the field of animal welfare science since its conception by Harding et al. (2004) (Baciadonna and McElligott, 2015; Bethell, 2015; Harding et al., 2004). Rather than specifically measuring anxiety or depression, the task purports to measure affective valence; the relative pleasantness or unpleasantness of the current affective state of the individual. To achieve this, the judgement bias task examines decision-making under ambiguity. Although there is some variation in methodology, the basic principles of the task outlined here are applicable to all judgement bias studies. Individuals are first trained to associate the presentation of one reference cue (e.g. a high frequency tone) with a reward and presentation of another reference cue (e.g. a lower frequency tone) with a lower reward or punisher. Once training is complete, individuals are presented with one or a few untrained probe cues that are intermediate between the reference cues (e.g. medium frequency tones). Whether an individual responds to these ambiguous cues as though they signal the more positive or less positive outcome, which is either measured as latency to approach the cue or choice to execute or not execute the riskier action which could lead to either the favourable or less favourable outcome (i.e. not the safe action which leads to a null outcome), is considered to be indicative of their expectation of rewards or punishers, and hence affective state. Individuals that exhibit a decreased latency to approach the cue, or more frequently execute the riskier action (often deemed 'more optimistic' or 'less pessimistic'), particularly at the ambiguous cues given that affect is considered to exert the greatest effect on decision-making when there is uncertainty about an outcome (Mendl et al., 2009), are considered to be in a relatively more positive affective state.

The task is based on the empirical finding that humans experiencing anxiety and depression have a greater expectation of punishing events or reduced expectation of

rewarding events than clinically healthy humans (Johnson and Tversky, 1983; Nygren et al., 1996; Wright and Bower, 1992). To assess the extent to which judgement bias could measure subjective affective state in humans, the task has been back-translated to human subjects. Studies using the back-translated judgement bias task have demonstrated a correlation between judgement bias and measures of subjectively experienced affect, such as the State-Trait Anxiety Inventory (STAI), Visual Analogue Scale for Anxiety (VAS-A), and negative affect dimension of the Positive and Negative Affect Schedule (PANAS) (Anderson et al., 2012; Iigaya et al., 2016; Paul et al., 2011). The finding that judgement bias correlates with subjective reports of affective state in humans supports judgement bias as measure of affect, and hence the task appears to have strong construct validity.

Research has been conducted to assess how judgement bias is influenced by affect-altering drugs in non-human animals (see Table 2.1). Synthesis of these studies would allow conclusions to be drawn about the ability of the judgement bias task to measure pharmacologically induced positive (relatively more pleasant) or negative (relatively less pleasant) affective states, and hence elucidate the validity and reliability of judgement bias as a measure of affect in non-human animals. To this end, we conducted a systematic review and meta-analysis to assess whether pharmacological manipulations alter judgement bias and hence assess the predictive validity of the task. In addition to assessing whether there was an overall effect, we investigated whether the relationship between affect-altering drugs and judgement bias was moderated by factors relating to the drug and administration of the drug, such as the duration and timing of administration, dosage, and neurobiological target of the drug. The potential moderating effects of several task-related factors, such as the presented cue, species used, sex, reinforcement type, response type, and the outcome measure, were also investigated. While we predicted that the effects of judgement bias would be greatest at the ambiguous cues and would depend on dosage, we did not predict that the other moderators would influence the effect of the pharmacological manipulations on judgement bias.

2.2 Methods

2.2.1 Literature search

This study followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement (Moher et al., 2009; see Fig. 2.1). A literature search was first conducted on the 2nd November 2016 to identify all judgement bias studies; the research articles from this literature search were split into groups of those that used pharmacological manipulations (to be analysed here) and those that did not (to be analysed in a separate analysis by Nakagawa et al., *in prep.*). These meta-analyses were conducted separately as they assessed different research questions; here we specifically want to examine the ability of judgement bias to detect pharmacological manipulations proposed to alter affect and to better understand the factors

moderating this, but also due to the complexity the use of different drug doses adds the meta-analysis.

In addition to these articles, a literature search was conducted on the 13th November 2017 using Scopus and Web of Science to identify more recent research papers, and a further literature search was conducted on the 12th July 2019 using Scopus and Web of Science, as well as additional searches in other subject databases (including PsycINFO, PsycARTICLES, PsycBOOKS, PsycEXTRA, PsycTESTS, EMBASE and Medline), grey literature (using ProQuest Dissertation and Thesis Database, Google Data Search and Dimensions platform), and snowballing from reviews on the topic (cited and citing references collected) and from already included papers (citing references collected). Further details on the literature search, including the search-terms used, can be found in Appendix A.

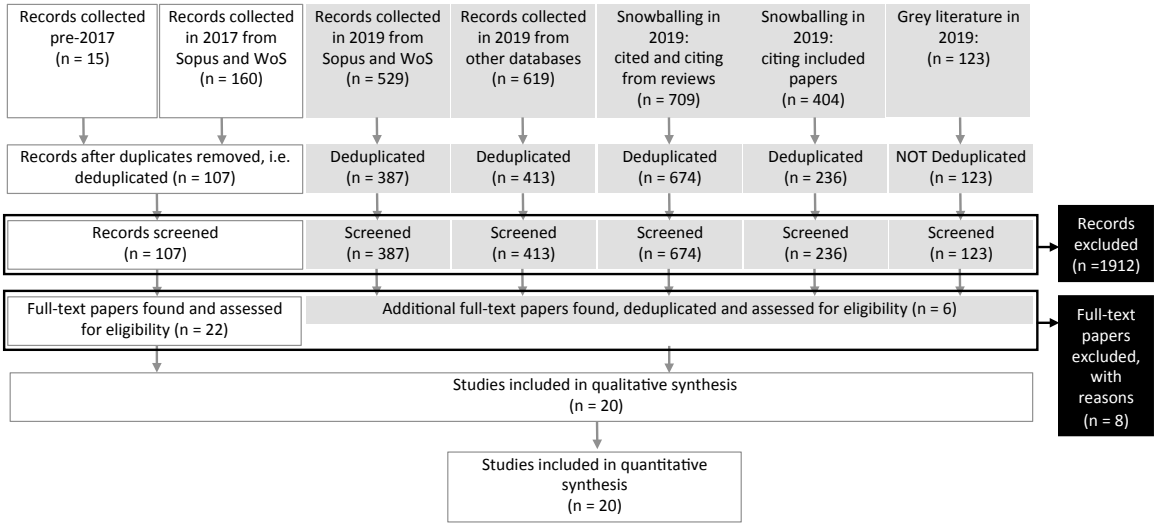


FIGURE 2.1: PRISMA Flow Diagram illustrating the number (n) of articles included at each stage of the literature review

2.2.2 Inclusion and exclusion criteria

Following removal of duplicates, the identified articles were first screened solely by their abstract. During this abstract-based selection, articles were selected for tentative inclusion in the meta-analysis if they were deemed to be an empirical study which compared judgement bias between at least one control group and at least one treatment group to whom an affect-altering drug had been administered. Additionally, to be included, these studies had to be conducted on vertebrate non-human animals. An affect-altering drug was classified as any substance that was considered to have antidepressant, depressant, anxiolytic, or anxiogenic effects. Twenty-eight articles met these inclusion criteria.

In the full-text screening, articles were selected on the basis that they had used a variant of Harding *et al*'s (2014) cognitive judgement bias task to compare judgement

bias between a group of individuals to whom a vehicle substance had been administered and at least one treatment group who had been given a drug to induce a positive or negative shift in their affective state (Harding et al., 2004). To be included, the outcome measure had to be either latency to approach the cue (e.g. a location in a test arena) on each trial, where approaching the presented cue had been associated with reward and hence shorter latencies would be interpreted as greater 'optimism', or the proportion of positive responses to each presented cue where a greater proportion would be interpreted as greater 'optimism', or an outcome measure that could be converted into either form. For example, if the article reported the proportion of negative responses to each presented cue or reported the percentage of positive responses to each presented cue, the extracted data would be subtracted from one or divided by 100 respectively. All included articles either reported the proportion or latency, but not both, and hence only one of these measures was extracted for each article. Two articles were excluded at the full-text selection stage following retraction by their authors, two were excluded for not meeting the inclusion criteria, and a further two were excluded as they were conference abstracts that duplicated data presented in a journal article which was included in the analysis (Table 2.1). In addition to these six exclusions, two authors did not provide the requested data and so data from their articles could not be included in the meta-analysis (Table 2.1). A total of 20 articles were included in the meta-analysis (Table 2.1).

TABLE 2.1: Articles included in the meta-analysis full-text screening and (where relevant) reason for exclusion

Article number	Status	Authors	Artide title	Journal	Year	Reason for exclusion
1	Included	Anderson, M. H., Munafo, M. R., Robinson, E. S. J.	Investigating the psychopharmacology of cognitive affective bias in rats using an affective tone discrimination task	Psychopharmacology	2013	NA
2	Included	Destrez, A., Deiss, V., Belzung, C., Lee, C., Boissy, A.	Does reduction of fearfulness tend to reduce pessimistic-like judgment in lambs?	Applied Animal Behaviour Science	2012	NA
3	Included	Doyle R. E., Hinch, G. N., Fisher, A. D., Boissy, A., Henshall, J. M., Lee, C.	Administration of serotonin inhibitor p-Chlorophenylalanine induces pessimistic-like judgement bias in sheep	Psychoneuroendocrinology	2011	NA
4	Included	Enkel, T., Gholizadeh, D., Von Bohlen Und Halbach, O., Sanchis-Segura, C., Hurlemann, R., Spanagel, R., Gass, P., Vollmayr, B.	Ambiguous-cue interpretation is biased under stress-and depression-like states in rats	Neuropsychopharmacology	2010	NA
5	Included	Golebiowska, G., Rygula, R.	Effects of acute dopaminergic and serotonergic manipulations in the ACI paradigm depend on the basal valence of cognitive judgement bias in rats	Behavioural Brain Research	2017	NA
6	Included	Hales, C. A., Robinson, E. S. J., Houghton, C. J.	Diffusion modelling reveals the decision making processes underlying negative judgement bias in rats	PLoS One	2016	NA
7	Included	Hales, C.A., Houghton, C.J., Robinson, E.S.J.	Behavioural and computational methods reveal differential effects for how delayed and rapid onset antidepressants effect decision making in rats	European Neuropsychopharmacology	2017	NA
8	Included	Hymel, K. A., Sufka, K. J.	Pharmacological reversal of cognitive bias in the chick	Neuropharmacology	2010	NA
9	Included	Iqsseve, O. S., Beard, A. P., Guy, J.H, Bateson, M.	Anxiety-depression model Elevated levels of fiio stress hormone, corticosterone, cause "pessimistic" judgement bias in broiler chickens	Scientific Reports	2017	NA
10	Included	Kis, A., Hernadi, A., Kanizsar, O., Gacsi, M., Top I, J.	Oxytocin induces positive expectations about ambivalent stimuli (cognitive bias) in dogs	Hormones and Behavior	2015	NA
11	Included	McGuire, M. C., Williams, K. L., Welling, L. L. M., York, J.	Cognitive bias in rats is not influenced by oxytocin	Frontiers in Psychology	2015	NA
12	Included	Rygula, R., Golebiowska, J., Kregiel, J., Holuj, M., Popik, P.	Acute administration of lithium, but not valproate, modulates cognitive judgment bias in rats	Psychopharmacology	2015	NA
13	Included	Rygula, R., Papciak, J., Popik, P.	The effects of acute pharmacological stimulation of the 5-HT, NA and DA systems on the cognitive judgement bias of rats in the ambiguous-cue interpretation paradigm	European Neuropsychopharmacology	2014	NA
14	Included	Rygula, R., Szczech, E., Kregiel, J., Golebiowska, J., Kubik, J., Popik, P.	Cognitive judgment bias in the psychostimulant-induced model of mania in rats	Psychopharmacology	2015	NA
15	Included	Rygula, R., Szczech, E., Papciak, J., Nikiforuk, A., Popik, P.	The effects of cocaine and mazindol on the cognitive judgement bias of rats in the ambiguous-cue interpretation paradigm	Behavioural Brain Research	2014	NA
16	Included	Sahin, C., Doostdar, N., Neill, J. C.	Towards the development of improved tests for negative symptoms of schizophrenia in a validated animal model	Behavioural Brain Research	2016	NA
17	Included	Stracke J., Ottena,W., Tiedtscherer A., Puppe, B., Duijnn, S.	Serotonin depletion induces pessimistic-like behavior in a cognitive bias paradigm in pigs	Physiology and Behavior	2017	NA
18	Included	Stracke, J., Otten, W., Tiedtscherer, A., Wirthahn, M., Metges, C.C., Puppe, B., Duijnn, S.	Dietary tryptophan supplementation and affective state in pigs	Journal of Veterinary Behaviour	2017	NA
19	Included	Verbeek, E., Ferguson, D., Lee, C.	Are hungry sheep more pessimistic? The effects of food restriction on cognitive bias and the involvement of ghrelin in its regulation	Physiology and Behavior	2014	NA
20	Included	Verbeek, E., Ferguson, D., Quinquet de Monjour, P., Lee, C.	Generating positive affective states in sheep: The influence of food rewards and opioid administration	Applied Animal Behaviour Science	2014	NA
21	Not included	Anderson, M.H., Munafo, M.R., Robinson E.S.J.	The effects of acute psychopharmacological treatments on cognitive affective bias in rats	European Neuropsychopharmacology	2012	Conference abstract that duplicates Anderson et al 2013
22	Not included	Hales C., Barthett J., Arban R., Hengerer B., Robinson E.	Targeted infusions with rapid acting antidepressants reveal a role for the prefrontal cortex in mediating affective biases and decision making	Brain and Neuroscience Advances	2019	Data not available
23	Not included	Karagiannis, C.I., Burman, O.H.P., Mills, D.S.	Dogs with separation-related problems show a "less pessimistic" cognitive bias during treatment with fluoxetine (Reconcile™) and a behaviour modification plan	BMC Veterinary Research	2015	Data not available
24	Not included	Kregiel J., Golebiowska J., Popik P., Rygula R.	Dopamine induces an optimism bias in rats-pharmacological proof for the translational validity of the ambiguous-cue interpretation test	Behavioural Brain Research	2016	Retracted by author
25	Not included	Kregiel, J., Malek, N., Popik, P., Starowicz, K., Rygula, R.	Anandamide mediates cognitive judgement bias in rats	Neuropharmacology	2016	Retracted by author
26	Not included	Neill J., Gaebel W., Wölwer W., Toeller V.	NMDA receptor antagonists in rodents, relevance to negative symptoms of schizophrenia: A translational link to humans	European Archives of Psychiatry and Clinical Neuroscience	2015	Not experimental research article
27	Not included	Phillips B.U., Dewan S., Nilsson S.R.O., Robbins T.W., Heath C.J., Saksida L.M., Bussey T.J., Alsio J.	Selective effects of 5-HT2C receptor modulation on performance of a novel valence-probe visual discrimination task and probabilistic reversal learning in mice	Psychopharmacology	2018	Didn't use a variant of Harding et al's task
28	Not included	Sahin, C., Podda, G., Grayson, B., Marsh, S., Articoglu, F., Neill, J.C.	The deficit in anticipatory motivation as a negative symptom of schizophrenia: Phencyclidine treated rats exhibit pessimism in an optimistic bias task	European Neuropsychopharmacology	2015	Conference abstract that duplicates Sahin et al 2016

2.2.3 Data extraction

We extracted the mean and standard deviation of either the latency to approach the presented cue, or proportion of positive responses to the presented cue, as well as the sample size (number of subjects), for every pharmacological treatment and control group for each cue from each article (JZ and VN extracted the data which were checked by VN and SN). Data in a graphical format were extracted using GraphClick 3.0.3 (http://www.arizona-software.ch/graphclick/) or WebPlotDigitizer 4.1 (http://automeris.io/WebPlotDigitizer). As we extracted mean values, we acknowledge that there may have been variation in how the authors incorporated non-responses into their calculation of latency mean values which we cannot control for. If multiple drug doses had been used, these variables (i.e. mean, standard deviation, and sample size) were extracted for each dosage, and similarly, if there were test sessions that varied the duration between administration and testing or number of days of chronic drug administration, these variables were extracted for each test session. Data collected from both a vehicle and treatment group prior to drug administration were not included as these data did not provide information about the effect of the pharmacological manipulation on judgement bias.

The extracted treatment and control group data were categorised according to whether the pharmacological manipulation was expected to induce either a more or less positively valenced affective state, which the judgement bias test is predicted to measure. If an anxiogenic or depressant substance had been administered, as determined by the hypotheses stated in the published article, the treatment group was categorised as the less positive group, and the vehicle group was categorised as the more positive group (i.e. a relatively negative judgement bias was predicted in the treatment group relative to the control group). If an anxiolytic or antidepressant drug had been administered, which was also determined by the hypotheses stated in the article, the treatment group was categorised as the more positive group and the vehicle group categorised as the less positive group (i.e. a relatively positive judgement bias was predicted in the treatment group relative to the control group). If no hypotheses were stated in the article, this categorisation was based on the description and pharmacodynamics of the substance as outlined on the DrugBank database (Wishart et al., 2017). Where multiple doses had been administered, higher doses of anxiolytic or anxiogenic drugs were categorised as more positive whereas higher doses of anxiogenic or depressant drugs were categorised as less positive. This was based on the widespread finding that drugs exert greater effects at higher doses (Shaheed et al., 2016; Thase, 1998).

Information about the article and authors, drug and drug administration, and methodology were also extracted (Table 2.1 and 2.2). These included; the article title, institute or university at which the research was conducted (extracted but not shown in Table 2.1 or 2.2), the name of the drug, the dosing duration (chronic - where drugs were administered repeatedly, acute - where the drug was administered immediately before testing, or chronic wash-out - the period after drug administration had stopped

following chronic administration), the time between administration and testing (acute studies only), the number of days since the first dose (chronic studies only), the dosage (in mg/kg), the neurobiological target of the drug, the pharmacological manipulation type (antidepressant/anxiolytic or depressant/anxiogenic), the species tested, and the outcome variable used (latency or proportion), cue (positive reference cue, midpoint probe cue, negative reference cue, and where included the near negative probe cue and near positive probe cue; not shown in Table 2.1 or 2.2, although number of probe cues given instead), sex of the experimental subjects (all male, all female, or both male and female), reinforcement type used for the reference cue (reward-punisher - where the positive reference cue was rewarded and negative reference cue punished; reward-null - where the positive reference cue was rewarded and negative reference cue was not rewarded; or reward-reward - where the positive reference cue was rewarded with a high reward and negative reference cue was reward with a low reward), response type which reflected whether both or only one of the reference cues required an approach response (go/no-go - where the positive reference cue required an active response and the negative reference cue required no response, or go/go - where both reference cues required an active response), the proportion of probe trials in relation to the total number of trials, and cue type (reference or probe). To ensure that dosage was comparable between substances and species, each drug dose within a species was standardized by dividing the dosage (in mg/kg) by the standard deviation of all doses administered within each drug for each species.

TABLE 2.2: Information extracted from each article included in the systematic review and meta-analysis

Article number	Drug	Pharmacological target	Doses	Dosing frequency	Time between administration and testing	Number of administration days prior to testing	Number of administration days between final treatment and testing	Manipulation	Species	Sex	Response type	Reinforcement type	Outcome variable	Number of probe cues	Proportion of probe cues
1	diazepam	GABAergic system	0, 0.3, 1	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.3
	fluoxetine	serotonergic system	0, 1	chronic	NA	1,4,8,11,15,18	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.3
	fluoxetine	serotonergic system	0, 1	chronic (wash-out)	NA	NA	5,7	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.3
	fluoxetine	serotonergic system	0, 0.3, 1, 3	acute	60	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	3	0.3
	reboxetine	adrenergic system	0, 0.1, 0.3, 1	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	3	0.3
2	diazepam	GABAergic system	0, 0.1	acute	10,180	NA	NA	antidepressant/anxiolytic	sheep	female	go/no-go	reward/punisher	latency	3	0.6
3	p-Chlorophenylalanine	serotonergic system	0, 40	chronic	NA	3,5	NA	depressant/anxiogenic	sheep	female	go/no-go	reward/punisher	proportion	3	0.6
	p-Chlorophenylalanine	serotonergic system	0, 40	chronic (wash-out)	NA	NA	5	depressant/anxiogenic	sheep	female	go/no-go	reward/punisher	proportion	3	0.6
4	corticosterone-HBC complex + reboxetine	multiple	0, 0.5 (cort) + 15 (rbx)	acute	30 (cort) + 60 (rbx)	NA	NA	depressant/anxiogenic	rat	male	go/go	reward/punisher	proportion	3	0.25
5	escitalopram	serotonergic system	0, 0.5, 1, 2	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.2
	haloperidol	dopaminergic system	0, 0.01, 0.02 0.05	acute	30	NA	NA	depressant/anxiogenic	rat	male	go/go	reward/punisher	proportion	1	0.2
	l-dopa	dopaminergic system	0, 2, 4, 8	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.2
6	FG7142	GABAergic system	0, 3.0, 5.0	acute	30	NA	NA	depressant/anxiogenic	rat	male	go/go	reward/reward	proportion	1	0.3
7	fluoxetine	serotonergic system	0, 0.3, 1	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/reward	proportion	1	0.3
	fluoxetine	serotonergic system	0, 1	chronic		1,4,8,11,15,18	NA	antidepressant/anxiolytic	rat	male	go/go	reward/reward	proportion	1	0.3
	fluoxetine	serotonergic system	0, 1	chronic (wash-out)	NA	4,7	NA	antidepressant/anxiolytic	rat	male	go/go	reward/reward	proportion	1	0.3
	reboxetine	adrenergic system	0, 0.3, 1	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/reward	proportion	1	0.3
	venlafaxine	multiple	0, 1, 3	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/reward	proportion	1	0.3
	ketamine	multiple	0, 0.3, 1, 3	acute	60	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/reward	proportion	1	0.3
	phencyclidine	multiple	0, 0.3, 1, 3	acute	40	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/reward	proportion	1	0.3
	amphetamine	multiple	0, 0.1, 0.3	acute	15	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/reward	proportion	1	0.3
	cocaine	multiple	0, 0.3, 1, 3	acute	10	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/reward	proportion	1	0.3
8	imipramine	multiple	0, 15	acute	15	NA	NA	antidepressant/anxiolytic	chicken	male	go/no-go	reward/punisher	latency	2	0.5
	clonidine	adrenergic system	0, 0.1	acute	15	NA	NA	antidepressant/anxiolytic	chicken	male	go/no-go	reward/punisher	latency	2	0.5
9	corticosterone	glucocorticoid system	0, 4	chronic	NA	3,4,5	NA	depressant/anxiogenic	chicken	female	go/no-go	reward/punisher	latency	3	0.3
10	oxytocin	oxytocin system	NA	acute	40	NA	NA	antidepressant/anxiolytic	dog	mixed	go/no-go	reward/null	latency	1	0.3
11	oxytocin	oxytocin system	0, 0.001	acute	5	NA	NA	antidepressant/anxiolytic	rat	male	go/no-go	reward/punisher	latency	1	0.25
12	lithium chloride	multiple	0, 10, 50, 100	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.2
	valproic acid	GABAergic system	0, 100, 200, 400	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.2
13	citalopram	serotonergic System	0, 1, 5, 10	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.2
	d-amphetamine	multiple	0, 0.1, 0.5, 1	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.2
	desipramine	multiple	0, 1, 2, 5	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.2
14	cocaine	multiple	0, 10	chronic	NA	14	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.2
	d-amphetamine	multiple	0, 2	chronic	NA	14	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.2
15	cocaine	multiple	0, 1, 2, 5	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.2
	mazindol	multiple	0, 0.5, 1 2	acute	30	NA	NA	antidepressant/anxiolytic	rat	male	go/go	reward/punisher	proportion	1	0.2
16	phencyclidine	multiple	0, 2	chronic (wash-out)	NA	NA	8, 9, 10, 11, 12	depressant/anxiogenic	rat	female	go/go	reward/reward	proportion	1	0.5
17	p-Chlorophenylalanine	serotonergic system	0, 50	chronic (wash-out)	NA	NA	1,2,3,8,9,10	depressant/anxiogenic	pig	female	go/no-go	reward/punisher	latency	3	0.14
18	tryptophan	serotonergic system	NA	chronic	NA	6,7,8,13,14,15	NA	antidepressant/anxiolytic	pig	female	go/no-go	reward/punisher	latency	3	0.14
19	ghrelin	multiple	0, 0.007	acute	10	NA	NA	depressant/anxiogenic	sheep	female	go/no-go	reward/punisher	proportion	3	0.6
20	morphine	opioid system	0, 1	acute	10	NA	NA	antidepressant/anxiolytic	sheep	female	go/no-go	reward/punisher	latency	3	0.6
	naloxone	opioid system	0, 2	acute	10	NA	NA	depressant/anxiogenic	sheep	female	go/no-go	reward/punisher	latency	3	0.6

2.2.4 Effect size and sampling variance calculation

As latency data are bounded at zero and proportion data are bounded between zero and one data obtained from the judgement bias task do not follow a Gaussian or normal distribution. The delta method (Taylor approximation) was used to adjust the extracted mean ($\bar{x}$) and (sampling) variance (sd^2) prior to calculating the effect size to account for the non-normality of the raw data (Nakagawa et al., 2017a). For extracted latency data, which were assumed to follow a log-normal distribution, this adjustment was calculated via the log transformation as:

$$\overline{ln(x)} = \ln(\bar{x}) - ln\left(\sqrt{1 + \frac{sd^2}{\bar{x}^2}}\right) \tag{2.1}$$

$$sd^2_{\ln} = \ln\left(1 + \frac{sd^2}{\bar{x}^2}\right) \tag{2.2}$$

In this case, the transformed sampling variance is exact and not an approximation.

For extracted proportion data, which were assumed to follow a binomial distribution, this adjustment was calculated via the logit transformation as (Nakagawa and Schielzeth, 2010):

$$\overline{\text{logit}(x)} = \text{logit}(\bar{x}) + \frac{sd^2}{2}\left(\frac{1}{(1-\bar{x})^2} - \frac{1}{\bar{x}^2}\right) \tag{2.3}$$

$$sd^2_{\text{logit}} = sd^2\left(\frac{1}{\bar{x}} + \frac{1}{1-\bar{x}}\right)^2 \tag{2.4}$$

Hedge's g (Hedges, 1981), a measure of effect size based standardized differences in means, was then calculated as the difference between the means of the relatively positive treatment (in which a relatively more positive affective state was expected, as outlined above) $\bar{x}_{+ve}$ and means of the relatively negative treatment (in which a relatively less positive affective state was expected, as outlined above) $\bar{x}_{-ve}$, divided by the pooled standard deviation, sd_{pool}, and then adjusted for biases arising from small sample sizes by factor J which depended on the sample size of the relatively positive n_{+ve} and relatively negative n_{-ve} groups:

$$\text{SMD} = \frac{\bar{x}_{+ve} - \bar{x}_{-ve}}{sd_{pool}} J \tag{2.5}$$

$$sd_{pool} = \sqrt{\frac{(n_{+ve}-1)sd^2_{+ve} + (n_{-ve}-1)sd^2_{-ve}}{n_{+ve} + n_{-ve} - 2}} \tag{2.6}$$

$$J = \left(1 - \frac{3}{4(n_{+ve} + n_{-ve}) - 9}\right) \tag{2.7}$$

For the latency data, Hedge's g was multiplied by minus one to account for a higher proportion being equivalent to a lower latency, in terms of judgement bias.

The sampling variance was calculated as follows:

$$se^2_{\text{SMD}} = \frac{n_{+ve} + n_{-ve}}{n_{+ve}n_{-ve}} + \frac{\text{SMD}^2}{2(n_{+ve} + n_{-ve})} \tag{2.8}$$

To account for shared controls, if one vehicle treatment group was compared to multiple drug treatment groups, an additional effect size and sampling variance was calculated based on a sample size for the vehicle group that had been divided by the number of treatment groups (Higgins and Green, 2011).

2.2.5 Meta-analysis and meta-regression models

The meta-analysis and meta-regression were conducted using the function, rma.mv from the R (R Core Team, 2017) package metafor (Viechtbauer, 2010); this function allowed us to fit multilevel meta-analytic and meta-regression models (Nakagawa et al., 2017b). All models included drug, institution at which the research was conducted, and effect ID (a unique ID given to each effect size) as random effects to account for the non-independence of effect sizes from studies conducted at the same institute or using the same drug (Van den Noortgate et al., 2013), and were fit using restricted maximum likelihood. The Knapp and Hartung adjustment was applied to all analyses (Knapp and Hartung, 2003). Initially, an intercept only model was fit to the effect sizes. A p-value for this model was obtained using a Wald-type test based on a t-distribution. Heterogeneity was assessed by calculating the I^2 values for each random effect in the model and an overall I^2 value for the model, following (Nakagawa and Santos, 2012), which is an extension of the original I^2 (Higgins and Thompson, 2002). The I^2 describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error, an I^2 value greater than 75% is considered to be high (Higgins and Thompson, 2002; Nakagawa and Santos, 2012).

Meta-regression was used to examine whether the following categorical and continuous moderators significantly contributed to variation between effect sizes: the dosing duration (chronic, acute, or chronic wash-out), the time between administration and testing (acute studies only), the number of days since the first dose (chronic studies only), the dosage differences between treatments from which the effect size was calculated, the neurobiological target of the drug, the manipulation type (positive or negative affect induction), the species tested, and the outcome variable used (latency or proportion), presented cue (positive reference cue, near-positive probe cue, midpoint probe cue, near-negative probe cue, negative reference cue), sex of the experimental subjects (all male, all female, or both male and female), reinforcement type (reward-punisher, reward-null, or reward-reward), response type (go/no-go, or go/go), cue type (reference or probe), and proportion of probe cues to reference cues in the test session. An omnibus test based on an F distribution, which examines the degree of variance explained by a moderator, was used to assess the significance of each moderator (Viechtbauer, 2010). To further investigate significant moderators, pairwise comparisons were made between the mean effect size for each level of the

moderator. A Wald-type test was used to assess the significance of these pairwise comparisons. Moderators which were significant in the meta-regression were subsequently included together in a full model and their influence on the effect sizes was re-assessed. To verify that the model of best fit included all moderators, Akaike's information criterion (AIC) was calculated for the full model and was compared to models where a moderator had been removed.

2.2.6 Subset analyses

As affect is hypothesised to provide information about the probability of each potential outcome of a decision, particularly when there is greater uncertainty about the potential outcomes, any treatment designed to influence affective state is expected to have the greatest influence on judgement bias at the ambiguous probe cues (see Fig. 2.2 for example of hypothesised data) (Mendl et al., 2009, 2010). There are also methodological and theoretical reasons as to why an effect may be observed at one cue and not others. For example, a cue may be too perceptually similar to either of the reference cues for there to be ambiguity about the outcome, or a potential punisher may be much more aversive than the reward is rewarding, to the extent that all animals will avoid probe cues that are similar to the negative reference cue. By considering all cues equally (including reference cues), the effect of an affective manipulation might be obscured, potentially leading to the false inference of no significant effect. To this end, we conducted an additional analysis on a subset of data that included only the effect sizes from the probe cue with the largest absolute effect size for each drug within an article. Additionally, we analysed a second subset of data that included only the effect sizes for the cue with the absolute largest effect size in the direction of the mean effect size for each drug within an article to avoid including outlying effects that might not necessarily reflect the influence of the affective manipulation. If only one probe cue was presented in a study, data from this probe cue were included in the subset data.

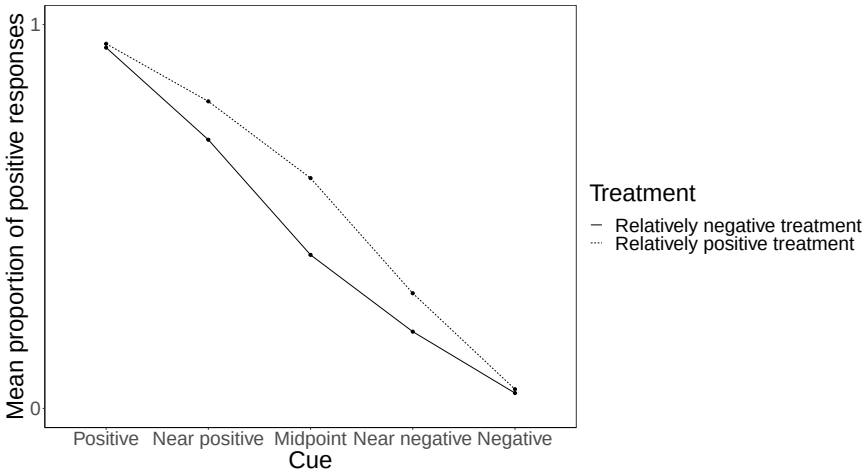


FIGURE 2.2: Example of hypothesised data from the judgement bias task with two treatments; one designed to induce a relatively positive affective state (relatively favourable treatment) and another designed to induce a relatively negative affective state (relatively unfavourable treatment). While the mean proportion of positive responses is almost identical at the positive and negative reference cue, a treatment difference is observed at the probe cues

2.2.7 Publication bias and sensitivity analysis

To assess the reliability of results across different analytical approaches and to check for a publication bias, the intercept-only and full meta-regression model were re-fit to the data under a Bayesian statistical framework using the R package MCMCglmm (Hadfield, 2010). The non-independence of effect sizes can also be accounted for using Bayesian methods. A parameter-expanded prior, allowing variance components to have different prior distributions, was used for both the random effect of drug and institution ID, while the prior variance for random effect of effect ID was fixed at one (i.e specified using the following R code):

```
prior=list(R=list(V=1, nu=0.002),
    G=list(G1=list(V=1, nu=1, alpha.mu=0, alpha.V=1000),
    G2=list(V=1, nu=1, alpha.mu=0, alpha.V=1000),
    G3=list(V=1, fix=1))
```

Model fitting had 110,000 iterations, 10,000 burn-in periods, and thinning by every 100, resulting in an effective sample size of 1000. The result of this intercept-only model was compared to our initial intercept-only model. The 'meta-analytic residuals' (*sensu* Nakagawa and Santos, 2012) from the full meta-regression model conducted in MCMCglmm were used to produce a funnel plot and run Egger's regression, which here regresses the meta-analytic residuals against precision where the precision is defined here as the inverse sampling error (Egger et al., 1997; Nakagawa and Santos, 2012), and hence checks for a publication bias. The meta-analytic residuals were calculated using the R function 'predict' in which a marginal effect of the sampling

error was specified. A significant result (i.e. $p<0.05$) using Egger's regression and an asymmetrical funnel plot would be indicative of a publication bias.

Additionally, the intercept-only meta-analysis was repeated but with the effect size and sampling variance that had been adjusted (via the sample size) for shared controls, to assess whether this altered the results.

2.3 Results

2.3.1 Data review

We extracted 557 effect sizes from 20 articles that had been published by authors based at 10 different institutions (see Table 2.1 and 2.2). Twenty-seven different drugs were used across these studies. The majority (328) of the effect sizes came from studies that had used drugs expected to induce a relatively positive affective state (anxiolytics or antidepressants, 12 articles), while the remainder used anxiogenic or depressant drugs (112 effect sizes, 9 articles). There were 408 effect sizes (14 articles) that came from studies using acute pharmacological manipulations, 97 effect sizes (6 articles) from studies using chronic pharmacological manipulations, and 52 effect sizes (5 articles) that came from the wash-out period of a chronic pharmacological manipulation. Most effect sizes came from studies using drugs that targeted the serotonergic system (198 effect sizes, 7 articles) while a high proportion of studies also used drugs that targeted a range of neurobiological systems (190 effect sizes, 9 articles) which included drugs such as cocaine and d-amphetamine which target the dopaminergic, serotoninergic, and adrenal systems. The remaining effect sizes were from experiments using drugs that specifically targeted GABAergic system (46 effect sizes, 4 articles), adrenergic system (43 effect sizes, 3 articles), dopaminergic system (36 effect sizes, 1 article), opioid system (20 effect sizes, 1 article), glucocorticoid system (15 effect sizes, 1 article), oxytocin system (9 effect sizes, 2 articles). Five different species were used across the studies: the most frequently used species according to the number of effect sizes was rat (418 effect sizes, 11 articles), followed by pig (60 effect sizes, 2 articles), sheep (50 effect sizes, 4 articles), chicken (23 effect sizes, 2 articles), and dog (6 effect sizes, 1 article). Proportion was more commonly used as the outcome measure (435 effect sizes, 12 articles) compared with latency (122 effect sizes, 8 articles). The majority of effect sizes came from studies using only male subjects (421 effect sizes, 11 articles), followed by only female subjects (130 effect sizes, 8 articles), and six effect sizes (1 articles) came from studies that used both male and female subjects. The most common reinforcement type was reward-punisher (420 effect sizes, 16 articles), followed by reward-reward reinforcement (131 effect sizes, 3 articles), and reward-null (6 effect sizes, 1 article). There were more effect sizes from studies using a 'go/go' design (415 effect sizes, 10 articles) compared with a 'go/no-go' design (142 effect sizes, 10 articles).

Across the articles from the acute studies, the average time between the administration of the drug and testing was 32.903±5.530 (mean±SE) minutes. The average

number of days between the start of the chronic drug treatment and testing was 9.000±1.074, and the average days the animal had been withdrawn from a drug when tested in the wash-out period was 6.938±0.824. The mean proportion of probe cues to reference cues used during a test session was 0.341±0.037. There were 11 articles that used more than one probe cue and three of these articles examined the effect of more than one drug. In total, there were 14 sets of effect sizes obtained from different articles using different drugs which used more than one probe cue. The probe cue with the greatest absolute effect size was the near-positive probe cue on nine occasions, the near-negative probe cue on four occasions, and the midpoint probe cue on one occasion. The probe cue with the greatest absolute effect size was also the presented cue with the greatest absolute effect size in the direction of the mean effect for all but one of the sets of effect sizes, where the near-positive probe cue had the greatest absolute effect sizes and the near-negative probe cue had the greatest absolute effect size in the direction of the mean effect.

2.3.2 Meta-analysis

Overall, considering all effect sizes equally, drugs expected to alter affective state did not significantly induce a judgement bias, although a small effect size (*sensu* Cohen, 1988: small=0.20, moderate=0.5, large=0.8) was observed (mean=0.239, 95% confidence interval or CI=-0.047-0.525, t_{556}=1.639, p=0.102). However, this needs to be interpreted in the context of the observed high total heterogeneity in the model, with an I^2 value of 89.535 (>75%=high, Higgins and Thompson, 2002), indicative of wide variation in the extent to which pharmacological manipulations alter judgement bias that warrants further examination. The between-effect-size effect (i.e., residuals; 42.378%) and the between-drug effect (i.e., which drug were used; 35.112%) explained a large percentage of this heterogeneity, while a smaller percentage of variability was due to institutional variation (12.044%). Heterogeneity between effect sizes was further explored through the meta-regression.

2.3.3 Subset analyses

However, as aforementioned, given the theoretical framework for judgement bias, we did not anticipate that effect sizes would be equal across all cues. Instead, we considered it likely that pharmacological manipulations would exert the greatest influence at only one of the probe cues, with proximity of this cue to the reference cues differing between studies as a result of different methodologies. Hence, subset analyses were conducted to assess the extent to which the pharmacological manipulations of affect altered judgement bias at, at least, one of the probe cues. Pharmacological manipulations of affect were found to have a significant small to moderate effect on judgement bias when the analysis was repeated on the subset data comprising only data from the ambiguous cue with the largest absolute effect size (mean=0.394, CI=0.017-0.770, t_{154}=2.067, p=0.040), and a significant small to moderate effect on the subset data

comprising only data from the ambiguous cue with the largest absolute effect size in the direction of the mean effect (mean=0.455, CI=0.061-0.849, t_{154}=2.279, p=0.024).

2.3.4 Meta-regression

The meta-regression revealed that several moderators significantly explained the observed heterogeneity among the extracted effect sizes, these moderators were: pharmacological manipulation type (Fig. 2.3: $F_{(1,555)}$=16.056, p<0.001), dosage (Fig. 2.3: $F_{(1,519)}$ =6.614, p=0.010), the reinforcement type (Fig. 2.3: $F_{(2,554)}$=3.653, p=0.027) and the cue type (Fig. 2.3: $F_{(1,555)}$=4.725, p=0.030). The presented cue (Fig. 2.3: $F_{(4,552)}$=2.002, p=0.093), and the neurobiological target of the drug were marginally non-significant moderators (Fig. 2.3: $F_{(8,548)}$=1.835, p=0.079). More specifically, pharmacological manipulations expected to induce a relatively negative affective state (either depressant or anxiogenic) had a greater effect on judgement bias than those expected to induce a relatively positive affective state (Table 2.3). Greater differences in dosage between the relatively positive and negative treatments were associated with smaller effect sizes. The greatest effect size was found when the reinforcement used for the reference cues was a high reward and low reward, compared with a reward and punisher. The effect of the pharmacological manipulation was greater at the probe cues compared with the reference cues (Table 2.3). The effect of the pharmacological manipulation was weaker at the positive reference cue compared to the midpoint probe cue, and near-positive probe cue, and tended to be weaker than the negative reference cue (Table 2.3). There was no difference in effect size at the positive reference cue compared with the near-negative reference cue (Table 2.3). The remaining moderators tested were not found to significantly explain variation in effect size. The effect of drugs targeting the adrenergic system differed significantly from all other drugs used apart from drugs targeting the opioid and oxytocin system (Table 2.3). Drugs targeting the adrenergic system had the opposite effect than expected; a negative judgement bias was induced when a positive judgement bias was hypothesised. Other moderators with non-significant effects included: species ($F_{(4,552)}$=0.835, p=0.503), dosing frequency ($F_{(2,554)}$=0.108, p=0.898), time since last dose ($F_{(1,406)}$ =0.467, p=0.495), number of days since first treatment ($F_{(1,95)}$=1.169, p=0.282), sex ($F_{(2,554)}$=0.328, p=0.720), response type ($F_{(1,555)}$=0.040, p=0.842), proportion of ambiguous cues to reference cues ($F_{(1,555)}$=1.531, p=0.217), and outcome measure ($F_{(1,555)}$=0.139, p=0.709).

TABLE 2.3: Pairwise comparison of each level of significant moderators from the meta-regression

Variable	**Model**	**Mean difference**	**CI lower bound**	**CI upper bound**	**p-value**
Cue	Midpoint - Positive	0.168	0.028	0.307	0.019
	Negative - Positive	0.119	-0.022	0.260	0.099
	Near Negative - Positive	0.181	-0.059	0.421	0.140
	Near Positive - Positive	0.253	0.015	0.492	0.037
	Midpoint - Near Positive	-0.086	-0.324	0.153	0.480
	Negative - Near Positive	-0.135	-0.375	0.105	0.270
	Near Negative - Near Positive	-0.073	-0.357	0.212	0.616
	Negative - Midpoint	-0.049	-0.190	0.091	0.493
	Near Negative - Midpoint	0.013	-0.227	0.253	0.914
	Negative - Near Negative	-0.062	-0.304	0.179	0.612
Cue type	Reference - Probe	-0.122	-0.232	-0.012	0.030
Manipulation Type	Negative - Positive	0.746	0.380	1.112	<0.001
Neurobiological Target	Serotoninergic - Adrenergic	0.895	0.118	1.672	0.024
	Dopaminergic - Adrenergic	1.152	0.254	2.049	0.012
	GABAergic - Adrenergic	1.362	0.540	2.184	0.001
	Glucocorticoid - Adrenergic	1.505	0.017	2.993	0.047
	Multiple - Adrenergic	1.056	0.346	1.766	0.004
	Opioid - Adrenergic	0.682	-0.365	1.729	0.201
	Oxytocin - Adrenergic	0.824	-0.676	2.324	0.281
	Serotoninergic - Multiple	-0.161	-0.622	0.300	0.493
	Dopaminergic - Multiple	0.096	-0.533	0.724	0.766
	GABAergic - Multiple	0.306	-0.244	0.855	0.276
	Glucocorticoid - Multiple	0.449	-0.914	1.812	0.518
	Opioid - Multiple	-0.374	-1.208	0.459	0.378
	Oxytocin - Multiple	-0.232	-1.608	1.144	0.741
	Dopaminergic - Serotoninergic	0.256	-0.435	0.948	0.467
	GABAergic - Serotoninergic	0.466	-0.156	1.088	0.141
	Glucocorticoid - Serotoninergic	0.610	-0.777	1.996	0.388
	Opioid - Serotoninergic	-0.214	-1.035	0.608	0.610
	Oxytocin - Serotoninergic	-0.071	-1.471	1.328	0.920
	GABAergic - Dopaminergic	0.210	-0.542	0.962	0.584
	Glucocorticoid - Dopaminergic	0.353	-1.114	1.821	0.636
	Opioid - Dopaminergic	-0.470	-1.469	0.529	0.356
	Oxytocin - Dopaminergic	-0.328	-1.807	1.152	0.664
	Glucocorticoid - GABAergic	0.143	-1.283	1.570	0.843
	Opioid - GABAergic	-0.680	-1.626	0.266	0.159
	Oxytocin - GABAergic	-0.538	-1.976	0.901	0.463
	Opioid - Glucocorticoid	-0.823	-2.355	0.708	0.292
	Oxytocin - Glucocorticoid	-0.681	-2.507	1.145	0.464
	Oxytocin - Opioid	0.142	-1.401	1.686	0.856
Reinforcement Type	Reward/Punisher - Reward/Null	0.150	-0.702	1.002	0.730
	Reward/Reward - Reward/Null	0.487	-0.395	1.369	0.279
	Reward/Reward - Reward/Punisher	0.337	0.089	0.585	0.730

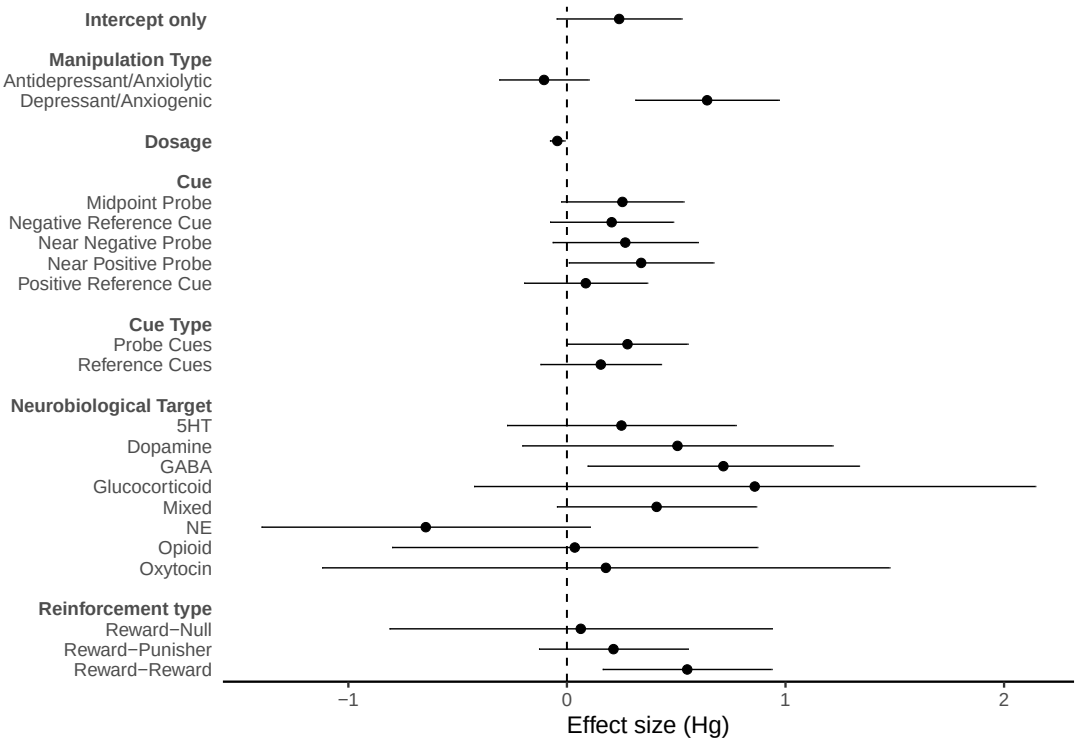


FIGURE 2.3: Forest plot with a meta-analytic mean (intercept-only model) and significant moderators from univariate meta-regression models. Each point represents the mean effect size for each moderator and error bars represent the 95% confidence interval

The best fitting model included cue (i.e. positive reference cue, midpoint probe cue, negative reference cue, near negative probe cue and near positive probe cue) instead of cue type (i.e. reference or probe; ΔAIC (i.e. difference in AIC values between models)=0.437), and all significant moderators identified in the univariate meta-regression. Removal of the neurobiological target of the drug (ΔAIC=13.511), dosage (ΔAIC=6.705), cue (ΔAIC=3.583), reinforcement type (ΔAIC=7.573), and manipulation type (ΔAIC=6.465) resulted in a poorer fit according to the AIC values. The best fitting model had a marginal R^2 value (*sensu* (Nakagawa and Schielzeth, 2013)) of 72.844%. In this model, the difference between effect sizes where a relatively positive compared with relatively negative affective state had been induced was significant, with a moderate effect size (Δmean=0.582, CI=0.054-1.110 t_{506}=2.164, p=0.031). Effect sizes from drugs targeting the adrenergic system were overall in the opposite direction to expected and there was a large and significant difference in effect sizes between adrenergic system targeting drugs and multiple system targeting drugs (Δmean=0.852, CI=0.043-1.661, t_{507}=2.069,p=0.039) and GABAergic system targeting drugs (Δmean=1.299, CI=0.369-2.228, t_{507}=2.746, p=0.006), and a large but marginally non-significant difference in effect sizes between adrenergic system targeting drugs and serotonergic system targeting drugs

(Δmean=0.817, CI=-0.073-1.707, t_{507}=1.803, p=0.072), dopaminergic system targeting drugs (Δmean=0.936, CI=-0.083-1.956, t_{507}=1.804, p=0.072), glucocorticoid system targeting drugs (Δmean=1.451, CI=-0.250-3.151, t_{507}=1.676, p=0.094). There was a large but non-significant difference between the effect sizes of drugs targeting the adrenergic and oxytocin system (Δmean=0.961, CI=-0.894-2.815, t_{507}=1.018, p=0.309) and moderate but non-significant difference between the effect sizes of drugs targeting the adrenergic compared with opioid targeting drugs (Δmean=0.555, CI=-0.621-1.732, t_{507}=0.927, p=0.354). Effect sizes were significantly weaker at the positive reference cue compared with the midpoint probe cue (Δmean=0.163, CI=0.019-0.308, t_{507}=2.218, p=0.027), and near-positive probe cue (Δmean=0.289, CI=0.025-0.553, t_{507}=2.148, p=0.032). Effect sizes at the positive reference cue were not significantly different from effect sizes at the negative reference cue (Δmean=0.118, CI=-0.028-0.264, t_{507}=1.592, p=0.112) or at the near-negative probe cue (Δmean=0.206, CI=-0.060-0.472, t_{507}=1.519, p=0.129). There was a small and significant difference between effect sizes from studies using high and low rewards as the more and less favourable outcome, respectively, compared with studies which used rewards and punishers (Δmean=-0.366, CI=-0.626–0.106, t_{543}=-2.764 p=0.006), while there was a moderate but non-significant difference between studies that used high and low rewards and those that used a reward and null outcome (Δmean=-0.546, CI=-1.548-0.457, t_{543}=-1.070, p=0.285).

2.3.5 Study exclusion

As the initial analysis revealed that drugs targeting the adrenergic system had the opposite effect on judgement bias than hypothesised, which differed significantly from the majority of the other drugs not specifically targeting the adrenergic system, our assumption about the effect of these drugs on affective state was called into question. Moreover, there is conflicting evidence about the affect altering properties of adrenergic system targeting drugs in non-human animals (Inoue et al., 2006; Tanaka et al., 2000). Consequently, we made the post-hoc decision to re-analyse the data excluding effect sizes from studies using adrenergic system targeting drugs. Three studies had used adrenergic system targeting drugs; one study had used clonidine and the other two studies had used reboxetine. Both clonidine and reboxetine are considered to induce a relatively positive affective state. These studies accounted for 7.719% (43) of the effect sizes analysed.

2.3.6 Post-exclusion meta-analysis

Following the exclusion of these effect sizes, a moderate overall effect was observed; pharmacological manipulations of affective state were found to significantly influence judgement bias in the predicted direction (mean=0.400, CI=0.056-0.744, t_{513}=2.287, p=0.023). However, there again existed high heterogeneity (I^2=89.746%); with 38.362% attributable to between-effect-size effects, 20.732% to between-drug effects,

and 30.653% to institutional variation. The meta-analysis using both data subsets (using only one probe cue) revealed a significant and moderate overall effect of pharmacological manipulations of affect on judgement bias (absolute greatest probe cue effect sizes: mean=0.520, CI=0.116-0.924, t_{144}=2.543, p=0.012; and absolute greatest probe cue effect in direction of mean: mean=0.579, CI=0.157-1.001, t_{144}=2.711, p=0.008).

2.3.7 Post-exclusion meta-regression

While manipulation type (Fig. 2.4: $F_{(1,512)}$=15.700, p<0.001) and dose (Fig. 2.4: $F_{(1,476)}$=5.169, p=0.023), remained significant as moderators when studies using adrenergic system targeting drugs were excluded, the presented cue (Fig. 2.4: $F_{(4,509)}$=2.396, p=0.049) was now significant as opposed to marginally non-significant and drug target ($F_{(6,506)}$=0.578, p=0.748) cue type ($F_{(1,512)}$=2.594, p=0.108), and reinforcement type ($F_{(2,511)}$=0.144, p=0.866) were no longer significant. The model which included all three significant moderators provided a better fit than the models which excluded manipulation type (ΔAIC=13.465), cue (ΔAIC=2.299), and dose (ΔAIC=5.430). This full model had a marginal R^2 value of 61.008%.

The difference between effect sizes at the midpoint probe cue and positive reference cue was very small but significant (Δmean=0.154, CI=0.011-0.297, t_{471}=2.111, p=0.035, and the difference between effect sizes at the positive reference cue and near-positive probe cue was small but marginally non-significant (Δmean=0.245, CI=-0.035-0.525, t_{471}=1.716, p=0.087). Contrary to the previous analysis including adrenergic-targeting drugs, a very small but significant difference was found between effect sizes at the negative and positive reference cues, with greater effect sizes at the negative reference cue (Δmean=0.183, CI=0.039-0.327, t_{471}=2.491, p=0.013). The difference in effect sizes between the positive reference and near-negative probe cue remained non-significant (Δmean=0.192, CI=-0.089-0.474, t_{471}=1.341, p=0.181). Effect sizes were still observed to be greater when the anxiogenic or depressant drugs were used compared to the antidepressant or anxiolytic drugs with a moderate difference in effect sizes (Δmean=0.701, CI=0.348-1.055, t_{471}=3.897, p<0.001), and effect sizes remained significantly greater when there were smaller differences in dosage between the relatively positive and relatively negative treatment, although the effect was very small (mean=-0.0379, CI=-0.071–0.004, t_{471}=–2.223, p=0.027).

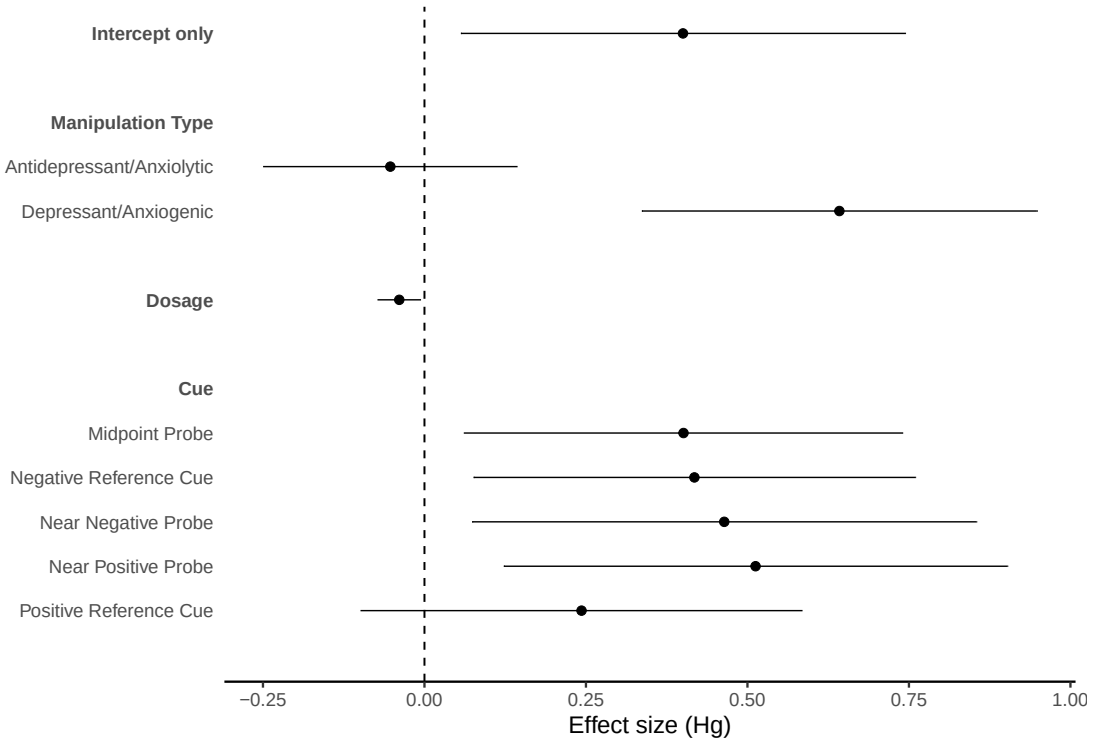


FIGURE 2.4: Forest plot with a meta-analytic mean (intercept-only model) and significant moderators from univariate meta-regression models following the exclusion of adrenergic system targeting drugs. Each point represents the mean effect size for each moderator and error bars represent the 95% confidence interval

2.3.8 Publication bias and sensitivity analysis

The results of the Bayesian meta-analysis were consistent with the results of our likelihood-based meta-analyses both prior to and following the removal of effect sizes from studies using drugs targeting the adrenergic system. The effect of the pharmacological manipulations on judgement bias was not significant prior to data exclusion (mean=0.242, 95% credible interval=-0.097-0.666, p=0.194), but a marginally non-significant overall effect emerged following data exclusion from studies using adrenergic system targeting drugs (mean=0.387, credible interval=0.020-0.864, p=0.056).

Visual inspection of the funnel plots produced from the meta-analytic residuals and raw effect sizes (Fig. 2.5) did not indicate that a publication bias was present, nor did the results of Egger's test on either the analysis prior to (t_{519}=-0.419, p=0.675) or following (t_{476}=0.568,p=0.570) the exclusion of data.

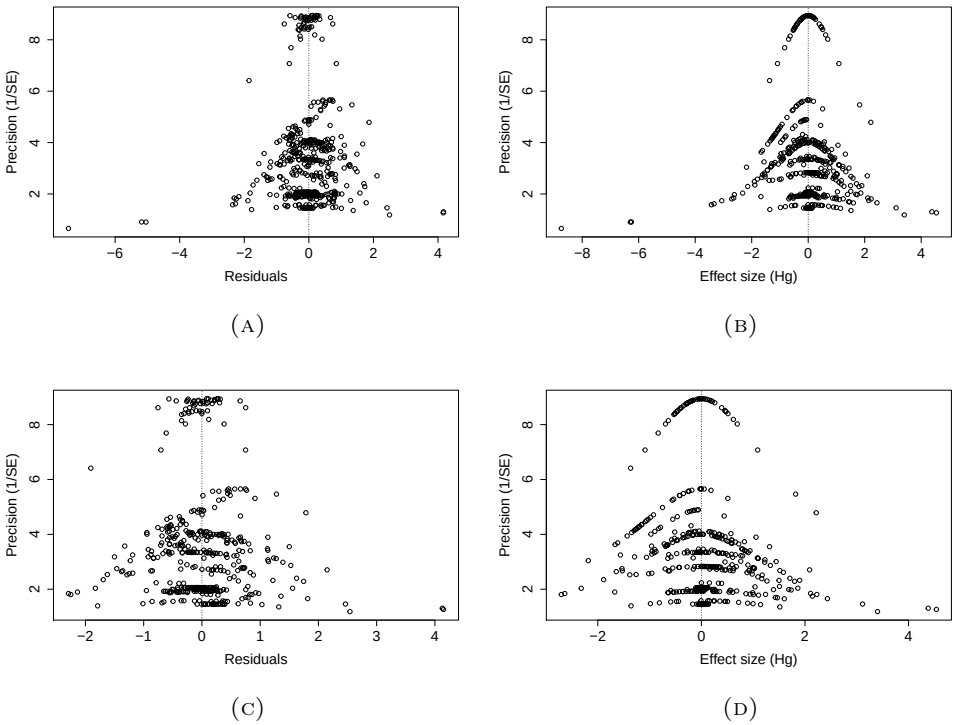


FIGURE 2.5: Funnel plots of a) the meta-analytic residual values (residuals + sampling errors) for the full meta-regression model prior to exclusion of effect sizes from studies using adrenergic system targeting drugs; b) the raw effect sizes and the inverse standard errors prior to exclusion of effect sizes from studies using adrenergic system targeting drugs; c) the meta-analytic residual values for the full meta-regression model following exclusion of effect sizes from studies using adrenergic system targeting drugs; d) the raw effect sizes and the inverse standard errors following exclusion of effect sizes from studies using adrenergic system targeting drugs

Re-analysis of the intercept-only model using the effect sizes and variances that had been adjusted for shared controls did not alter the results qualitatively. The result prior to data exclusion was statistically non-significant (mean=0.240, CI=-0.047-0.527, t_{522}=1.641,p=0.101) while following data exclusion was significant (mean=0.401, CI=0.057-0.745, t_{476}=2.291, p=0.022).

2.4 Discussion

Judgement bias is a relatively new and promising measure of animal affect that may provide a useful alternative to more common behavioural assays used to assess the efficacy of potential pharmacological treatments of mood disorders, such as the forced swim test. Empirical studies with human subjects have supported its construct validity (Anderson et al., 2012; Johnson and Tversky, 1983; Nygren et al., 1996; Paul et al., 2011; Wright and Bower, 1992). To examine its predictive validity, we conducted a systematic review and meta-analysis of studies investigating the effect of

affect-altering drugs on judgement bias in non-human animals. We analysed data from 20 published research articles which yielded 557 effect sizes.

There was high heterogeneity (>75%) between the effect sizes observed (Senior et al., 2016) indicating strong variability in the extent to which pharmacological manipulations of affective state alter judgement bias. The drug used accounted for some of this variability, as did the institution at which the research was conducted, yet a high proportion of heterogeneity was also attributed to variation within drug and institution. Our meta-regression further highlighted a number of factors which explained variation in effect sizes including the neurobiological drug target, manipulation type (whether the drug was hypothesised to induce a negative or positive affective state), dosage, cue, and cue type (reference or probe).

Initially, considering all effect sizes across all cues equally (including references cues), we found no significant overall effect of affect-altering drugs on judgement bias. However, because there are theoretical and empirical reasons for an effect being more likely at ambiguous cues as opposed to reference cues, and/or to occur at one ambiguous cue but not others (e.g. because the others may happen to be too perceptually similar to the reference cues, see Methods), considering all cues equally may obscure an effect of a treatment manipulation. Indeed, judgement bias studies often report effects that are observed at only a subset of (ambiguous) cues (e.g. Bethell and Koyama (2015)). Consequently, we also carried out analyses using subsets of data that included only (i) effect sizes for the probe cue with the largest absolute effect size; (ii) effect sizes for the cue with the absolute largest effect size in the direction of the mean effect size. These analyses revealed that the pharmacological manipulations altered judgement bias in the predicted direction.

The results of the meta-regression showed a clear moderating effect of the neurobiological drug target, particularly of drugs targeting the adrenergic system, whose effect differed significantly from the majority of other drugs used. A small to medium effect using data from all cues was found following the removal of data from studies targeting the adrenergic system, and a moderate effect was found when considering data from the subset analyses described above. Thus, this meta-analysis supports judgement bias as measure of affect in non-human animals, having demonstrated that pharmacological manipulations of affect overall alter judgement bias at the probe cues in the predicted direction.

The three excluded studies used either reboxetine, an antidepressant, or clonidine, which is used off-licence to treat anxiety disorders. Jointly, these drugs were found to exert an opposite effect on judgement bias; inducing a negative judgement bias when a positive judgement bias was predicted. Paradoxically, depression and anxiety are known side effects of these drugs (Center for Drug Evaluation and Research, 2019). Moreover, these studies both used an acute dose which may explain why their effects were not in the predicted direction. Both norepinephrine and cortisol increase in response to stress and acute dosing of drugs which simultaneously elevate levels of norepinephrine and cortisol have been shown to result in stress-like changes in the

neural response to negative stimuli in humans (Kukolja et al., 2008). It is therefore feasible that the acute delivery of adrenergic system targeting antidepressant drugs induced a relatively negative rather than a positive affective state which resulted in the relatively negative judgement bias observed. This potential explanation is further supported by studies that have observed anxiety-like states in rodents following the administration of similar adrenergic system targeting drugs (Inoue et al., 2006; Tanaka et al., 2000).

An alternative explanation could be related to another side effect of adrenergic agonists that has been documented in human and non-human animal subjects; sedation (Buerkle and Yaksh, 1998; Center for Drug Evaluation and Research, 2019) A sedated animal may not have been able to fully partake in the experiment or have been considerably slower to respond, leading to seemingly 'pessimistic' responses. This is perhaps further supported by the finding that clonidine and reboxetine led to an increased latency to respond to the positive reference cue, as described by the authors of the studies included in this meta-analysis. However, further studies would be required to reveal the extent to which the results from these two studies can be generalised to all adrenergic system targeting drugs.

Drugs inducing negative affect (e.g. depressants and anxiogenics), had a greater effect on measured judgement bias than drugs inducing positive affect (e.g. antidepressants and anxiolytics). This result may reflect an interaction between the drugs administered and affective states arising from the process of being tested, which may sometimes be negative in their own right (e.g. invasive administration of drugs, social isolation during testing, and potential delivery of an aversive decision outcome). These factors may have enhanced the effect of the negative affect inducing drugs, while dampening the effect of the positive affect inducing drugs. Notably, the potential negative affective state induced by testing may also explain the finding, prior to exclusion of adrenergic targeting drugs, that effect sizes were greater when only rewards were used as reinforcement, as opposed to both reward and punishers. Indeed, in humans there is evidence to suggest that some affect-altering recreational drugs intensify the affective state of an individual prior to consumption, or result in the exaggerated interpretation of emotional stimuli (Foisy et al., 2007; Parrott et al., 2011). With regards to the development of novel treatments for mood disorders such as depression and anxiety, this perhaps suggests that attention should be given to the potential effects of the testing procedure on affect, and that greater sample sizes may be required to provide sufficient power for an effect of the potential pharmacological treatment to be detected.

Another possible explanation for the moderating effect of manipulation type is that there are floor effects which limit the impact positive affect inducing drugs may have on judgement bias. There will be a physical limit to how quickly an animal can approach a cue, and the control animals may already be performing at or close to this limit, meaning that the animals that had been given the positive drug could not respond any quicker. However, this explanation will only be relevant to studies measuring approach

latency. Similarly, the smoke-detector principle states that individuals should be overly responsive to potential threats (Nesse, 2001); just as the cost of a smoke detector not detecting a fire is far greater than the cost of the smoke detector sounding an alarm when there is no fire, false positives are also optimal in the detection of predators. An individual may continue to appear relatively cautious even when in a more positive affective state because the cost of not avoiding punishers (i.e. potential death) is so high that it would be suboptimal for an individual to behave in a more risky manner (i.e. making the 'optimistic' response which could lead to a punisher, as opposed to the 'pessimistic' response which is the safe option) (Nesse, 2001; Nettle and Bateson, 2012).

Although this meta-analysis did not identify a difference in effect sizes between studies which analysed proportion or latency data, it is important to discuss these two outcome measures. Typically, both latency and proportion data can be collected from judgement bias studies and it is unclear what drives a researcher to select either measure. There are merits and disadvantages of using either measure; while latency contains more information than proportion data, in the sense that as a continuous variable it may identify variation that proportion data cannot, it may also be more subject to influences from other factors such as any effect of the drug on motor responses as outlined above or other cognitive biases such as attention biases (Mendl et al., 2009).

Greater effects were observed when there were relatively smaller differences in dosage between treatments. This is consistent with the inverted U-shaped dose-response function that is sometimes observed in drug studies (Calabrese and Baldwin, 2001a,b,c). This result may reflect that higher doses increase the probability of side effects which may interfere with task performance (Furukawa et al., 2002; Lazarou et al., 1998). Adverse effects that alter the motivation of the animal (e.g. reduced appetite), their consummatory behaviour (e.g. nausea), or psychomotor abilities (e.g. sedation) are likely to affect judgement bias. Such side effects are common to several affect-altering drugs (Center for Drug Evaluation and Research, 2019). It may therefore be sensible to take measures of activity or food consumption concurrent to the judgement bias task to assess the potential impact of side effects of drug manipulations.

The meta-analysis also found that the effects of the pharmacological manipulations on judgement bias were weakest at the positive reference cue, and that the effect of pharmacological manipulations was greater when only the probe cue with the greatest effect size within each drug and article were analysed. This reflects that pharmacological manipulations of affect exert a stronger influence on trials where there is ambiguity about the outcome of the trial compared with trials where the reward is certain. On presentation of the positive reference cue, there should be little ambiguity about the outcome, and it would be expected that the animal should make the response that allows them to obtain the reward on a high proportion of trials. The influence of any affective manipulation should be greatest when there is uncertainty about the outcome

as subjective probabilities of uncertain outcomes are thought to be more strongly informed by an individual's affective state (Mendl et al., 2010; Mendl and Paul, prep; Trimmer et al., 2013). Thus, this finding is consistent with the theoretical framework underlying judgement bias. However, given that a number of cognitive processes could lead to a shift in judgement bias at the probe but not the positive reference cue, this finding does not negate the possibility that cognitive processes other than probability estimation underlie the relationship between the pharmacological manipulations and decision-making on the judgement bias task.

It is unclear why the extracted effect sizes were not smaller at the negative reference cue following exclusion of effect sizes from studies that had used adrenergic-targeting drugs. The pharmacological manipulations were not expected to exert a similar effect at the negative reference cue compared with the probe cues, as there should be little uncertainty about the outcome when the reference cues are presented. Moreover, in studies in which multiple probe cues were presented, the pharmacological manipulations rarely exerted the greatest influence at midpoint cue, where there should be the greatest uncertainty about the outcome. This further suggests that pharmacologically induced affective states do not necessarily induce a greater judgement bias as uncertainty about the decision outcome increases. Both valuation and probability of decision outcomes are key components of decision-making; an individual might be more likely to make a risky or more 'optimistic' response if they considered the reward to more probable or punisher to be less probable or if they considered the reward to be more valuable or punisher to be less aversive (Mendl et al., 2009; Rangel et al., 2008). Speculatively, it is possible that the pharmacological manipulations altered the valuation of the punisher, hence altering responses to its presentation. Indeed, several neurobiological systems implicated in affect have also been found to underlie reward and punisher valuation. For example, the dopaminergic system has been associated with the incentive salience ('wanting') of rewards; the opioid system has been associated with the hedonic value ('liking') of rewards; while the serotonergic system has been associated with punisher avoidance (Berridge et al., 2009; Boureau and Dayan, 2011). A better understanding of the cognitive processes underlying judgement bias, which could be achieved through a battery of tests or computational modelling of judgement bias data, would be highly valuable. Similarly, it is possible that some findings reflect differences between the reward and punisher systems, especially given that the majority of studies reinforced the negative cue using a punisher, as opposed to null outcome or lower reward.

Our meta-analysis found no evidence to indicate that the species used, the dosing frequency, the time since last dose in acute studies, the number of days since first treatment in chronic studies, the outcome variable used, the biological sex of the individuals studied, the reinforcement type, or response type had moderating effects on the influence of pharmacological manipulations of affect on judgement bias. While this might reflect that there is insufficient power to detect an effect, it might indicate

that judgement bias is robust to variation in methodology and across species. Interestingly, despite being one of the most commonly used non-human animal species in research, none of the studies included in this meta-analysis used mice (National Research Council, 1988; UK Home Office, 2017). As judgement bias tasks have been successfully conducted in mice (Hintze et al., 2018; Krakenberg et al., 2019), we consider that it would be highly worthwhile to examine the extent to which pharmacological manipulations alter judgement bias in mice. We found no evidence to suggest a publication bias.

Future studies should attempt to account for the potential side effects of pharmacological manipulations. Observing behaviour following drug administration, for example activity levels and food and water consumption, may help to highlight potential adverse effects. The majority of effect sizes extracted in this meta-analysis were from studies using serotonergic system targeting drugs. While this is unsurprising given that commonly prescribed antidepressants target the serotonin system (Mars et al., 2017), mood disorders are associated with dysfunction in several neurological systems and further investigation of the influence of pharmacologically-induced changes in the activity of these systems may be beneficial (Joëls and Baram, 2009; Nemeroff and Owens, 2002; Nestler et al., 2002). This meta-analysis has highlighted that multiple probe cues may be preferable in future studies. Pharmacological manipulations of affective state do not necessarily exert the strongest influence of judgement bias at the most ambiguous cue, as found in this meta-analysis, and using multiple cues would allow a more comprehensive assessment of the effect of the manipulation. Finally, it would be worthwhile to assess the efficacy of judgement bias as a measure of pharmacological manipulations of affect in mice.

2.5 Conclusions (see Fig. 2.6)

This meta-analysis has provided evidence that judgement bias has predictive validity as a measure of the affective impact of pharmacological manipulations. However, a key issue is the potential interference of drug side effects on judgement bias. In particular, the contrary effect of adrenergic-targeting affect-altering drugs and the greater effect of drugs on judgement bias at lower doses, may be attributed to side effects or to the complex nature of adrenergic drug effects. The effect of drugs hypothesised to induce a negative affective state was greater than the effect of drugs hypothesised to induce a positive affective state, and therefore larger sample sizes may be required when testing the efficacy of potential pharmacological treatments for mood disorders. However, if consideration is given to these potential shortcomings, the judgement bias task for which there is evidence of construct validity and now of predictive validity, appears to be a viable measure of pharmacologically-induced affect in non-human animals.

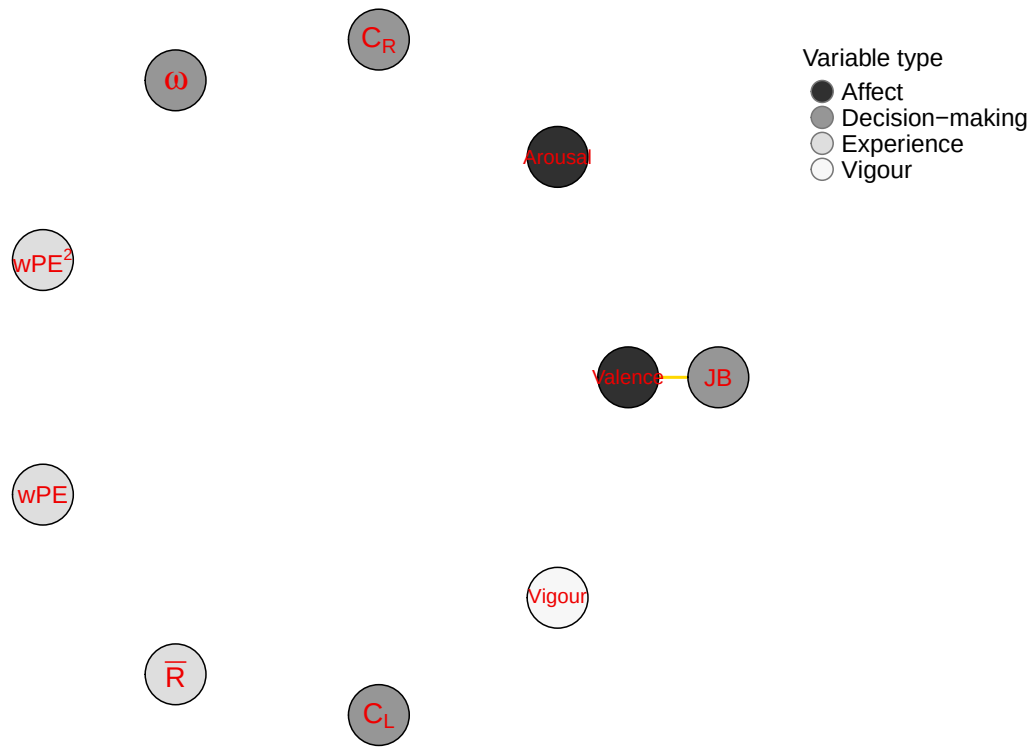


FIGURE 2.6: Diagrammatic summary of the results of Chapter 2: the nodes represent variables relating to affect, decision-making (where C_L denotes loss/punisher sensitivity, C_R denotes reward sensitivity, ω denotes the prior belief about outcomes, and JB denotes judgement bias), experience (where wPE denotes the weighted reward prediction error, $w\text{PE}^2$ denotes the squared weighted reward prediction error, and $\bar{R}$ denotes the average earning/reward rate), and vigour, the lines between nodes represent observed associations between those variables.

Chapter 3

How do fluctuating monetary rewards and losses influence human affect and judgement bias?

Chapter summary: An individual's ongoing experience of rewards and punishers is considered to be a fundamental determinant of both affect and decision-making. The judgement bias task uses a series of one-shot decisions to measure affect, but little consideration has been given to the influence that the outcomes of these decisions might exert on subsequent affect and decision-making. To better understand how affect and decision-making vary across a judgement bias test session in accordance with experience, which ultimately may inform our understanding of the adaptive function of affect, we examined the relationship between reward and punisher experience, affective state, and decision-making in humans. To do this, we conducted a judgement bias study in which either the offered monetary rewards or threatened monetary losses fluctuated systematically, and asked participants to regularly report both their affective valence and arousal. We also developed a novel computational model for the analysis of judgement bias choice data. We found that participants reported more positive affective valence and were faster to initiate trials in the task when the average earning rate was higher, although the average earning rate did not influence judgement bias. However, participants were more likely to opt for the risky choice when recent outcomes had been more predictable, which operated via loss sensitivity. Moreover, the extent to which the predictability of recent outcomes modulated judgement bias was significantly associated with affective valence; greater predictability-dependent modulation of judgement bias was associated with more positive valence. Thus, reward and punisher experience,

specifically the average earning rate and unpredictability of outcomes, influences affect and decision-making, which may suggest that these variables provide important information about environmental conditions.

3.1 Introduction

There is a traditional gulf in the field of decision-making between one-shot tasks, in which each trial is independent of the others, and ongoing tasks, in which subjects are supposed to apply what they learn from the consequences of their actions in earlier trials to later trials. Signal detection theory (Green et al., 1966) is the paradigm for the former, and reward-based reinforcement learning (Sutton and Barto, 1998) for the latter. In reality, the distinction is blurred – adaptation paradigms use repeated stimulus presentations to 'set' the state for one-shot psychophysical examinations (Clifford, 2002); subjects perform exquisite Bayesian learning in stop signal reaction time tasks, despite being instructed not to (Ma and Yu, 2015).

Here, we exploit the judgement bias task (Harding et al., 2004) as a paradigmatic example of this blurring. At the heart of the judgement bias task is a series of apparently independent one-shot psychophysical choices. However, decision-making on each trial, particularly trials with perceptually ambiguous stimuli, is considered to be indicative of a longer timescale notion of affective state (i.e. 'mood'), with greater risk-aversion typically associated with more positively valenced affect which should reflect approximate sufficient statistics about the characteristics of the interaction between an individual and their environment (Eldar et al., 2016; Mendl et al., 2010; Trimmer et al., 2011). These statistics could encompass such things as the present provision of rewards and punishers or surprise at the past provision of these outcomes, and could affect decision-making in various ways, for instance changing the prior expectations of subjects, or their sensitivities to reward and punisher (Eldar et al., 2016; Mendl et al., 2009, 2010; Nettle, 2008; Trimmer et al., 2011). To illustrate this, consider an individual in a reward-barren environment. Such an environment has been proposed to induce a depression-like state which is associated with a reduced reward valuation ('anhedonia') and a reduced expectation of future rewards, and hence should result in more risk-averse decision-making in the judgement bias task (Eldar et al., 2016; Mendl et al., 2009, 2010; Nettle, 2008; Trimmer et al., 2011). However, there are significant gaps in our knowledge about the bi-directional relationship between affect and choice in the judgement bias task: what characteristics of past decisions and outcomes influence affective state, and also how they influence present choices, either directly, or indirectly by manipulating affective state.

We aimed to examine this relationship by considering the impact of ongoing experience within the judgement bias task, in which either reward or loss magnitude is systematically varied, on psychophysical decision-making and self-reported affect in human participants. This systematic variation of reward or loss was inspired by

Beierholm et al. (2013) and allowed reward and loss sensitivity to be more easily disentangled through computational modelling (see Fig 3.1 for demonstration). The fluctuating offered reward or threatened loss also allowed greater variation in the values of the average earning rate and reward prediction error across trials so that the effects of these variables on decision-making would be more easily detected.

FIGURE 3.1: The probability of making the 'stay' response, calculated using a softmax function and assuming a 0.5 probability of each outcome, on three trials (n, n+1, n+2) with a varying offered reward and static threatened loss and different scaling (1.5×reward, no scaling, 1.5×loss) of the rewards and loss (i.e. reward and loss sensitivity). The probabilities of 'staying' across trials differs depending whether the reward or loss is scaled, allowing reward and loss sensitivity to be disentangled

A detailed computational model of a go/no-go version of the judgement bias task allowed us to investigate the cognitive mechanisms underlying the relationship between reward and punisher experience, affect, and decision-making. While decision-making (or latency) is the typical outcome measure in judgement bias tasks, the modelling approach used in this study sought to identify underlying, and perhaps more fundamental, variables that link affect, experience, and decision-making in the judgement

bias task. Here, we modelled decision-making in the judgement bias task under the framework of a partially observable Markov decision process (POMDP). The model considers that, on each trial, participants transition through a two-dimensional state space: one dimension represents their (discretised) belief about the presented stimulus which is informed by their observations, and the other dimension represents the (discretised) time elapsed on the trial. The probability that a participant executes a 'go' or 'no-go' action will depend on their transitions through this state space and subjective value of occupying each state. This is determined by a number of parameters, including those characterising task performance, such as a slope and a lapse parameter (Wichmann and Hill, 2001), and those characterising reward and punisher experience, such as those mapping the influence of the average earning rate, prediction error, and predictability to prior beliefs about the stimulus and reward and loss sensitivity. Since we expect predictions, prediction errors and surprise to play an important role in affect given that unexpected rewards have been proposed to induce positive affect (Eldar et al., 2016; Rutledge et al., 2014), we needed a robust quantification of the predictions that our subjects are making, which we derived from the computational model. We also examined the vigour with which participants engage in the task, as measured by latency to initiate trials, since this has been another prominent correlate of longer-term aspects of decision-making such as the average reward rate, with several studies demonstrating that vigour increases concurrently with the average rate of rewards or losses (Griffiths and Beierholm, 2017; Guitart-Masip et al., 2011; Niv et al., 2007).

We hypothesised that more positive prediction errors and a greater average earning rate (differing from the average reward rate in that it also encodes punishing and null outcomes) would reflect that the test environment is both relatively and absolutely favourable, thereby promoting activity and risk-seeking behaviour and hence a bias towards the risky choice on the judgement bias task, faster trial initiation, and would be associated with a positively valenced high arousal affective state (Eldar et al., 2016; Guitart-Masip et al., 2011; Mendl et al., 2010; Fig. 3.2). Environments that are unpredictable have also been proposed to induce negative affect (Clark et al., 2018; Fig. 3.2)), hence we hypothesised that we might also observe more negative affect and risk-averse decision-making following less predictable outcomes. We did not expect differences in the direction of these effects depending on whether the reward or loss was varied, although given that losses can be more salient than gains, consider that stronger effects might be observed when losses were varied (Kahneman and Tversky, 1979; Kermer et al., 2006; Taylor, 1991).

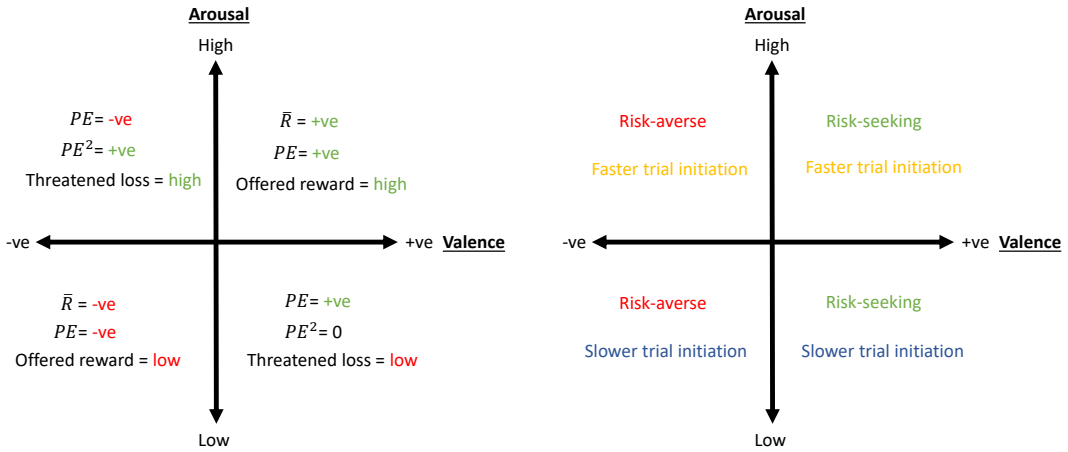


FIGURE 3.2: Diagrammatic summary of hypotheses; (A) different aspects of reward and punisher experience (i.e. the prediction error - PE; squared prediction error - PE^2; average earning rate - $\bar{R}$; and the offered reward and threatened loss) would modulate affective valence and arousal; (B) which in turn would alter risk-aversion and vigour with which trials were initiated

3.2 Methods

3.2.1 Participants

Thirty-nine students (34 female, 5 male, mean age ±SE=24.590±0.457) from the University of Bristol participated in the study and were paid £5 per session for their participation plus a performance-dependent bonus. Participants provided written, informed consent, and the study was approved by the Faculty of Science Research Ethics Committee at the University of Bristol.

3.2.2 Procedure

Twenty participants undertook the experiment in the context of obtaining variable rewards (fluctuating reward condition), and nineteen in the context of avoiding variable losses (fluctuating loss condition). The task was written in MATLAB (MathWorks, Natwick, MA, USA) using the PsychToolBox extensions.

3.2.3 Judgement bias task (see Fig. 3.4)

This task provided a human version of the rat judgement bias task described by Jones et al. (2018). Participants were required to make go/no-go responses to classify the direction of motion of random dot kinematograms (RDK) as leftwards or rightwards, with one direction being associated with a monetary reward and the other with a monetary loss. To initiate each trial, participants were required to press and hold the 'enter' key, this led to a fixation cross being displayed for 500ms, followed by a RDK displayed for 2000ms. The true direction of motion of the RDK was always either

leftwards or rightwards, and the proportion of dots moving coherently was varied to alter the difficulty with which the motion could be classified by the participant as leftwards or rightwards. Half of the participants were told that when the direction of motion was rightwards (threatened loss trials), they must release the 'enter' key ('go') prior to two seconds to avoid a loss (see below), and when the motion was leftwards (offered reward trials) they must continue to press the 'enter' key ('stay') for two seconds to obtain a monetary reward (see below) following an on-screen prompt to 'Please press and hold the enter key', while the other half of the participants were told the obverse (i.e. leftwards=threatened loss; rightwards=offered reward). The duration of the RDK display was 2000ms regardless of choice. In all blocks of trials, the true direction of motion was leftwards on half of the trials and rightwards on the remaining half, and each coherence level used within each block occurred an equal number of times within each direction of motion. The order of trials was randomised.

Participants completed three training blocks consisting of 24, 60, and 24 trials respectively, followed by one test block of 180 trials. In the first practice block, the direction of motion of the dots were unambiguous, with a coherence level of 0.32. In the second block, the direction of motion of the dots was difficult for the participants to determine (i.e. ambiguous) on a third of trials with a coherence level of 0.04, and unambiguous on two thirds of trials with coherence levels of 0.32 or 0.16. Stimuli in the third training and also the test block had a coherence level of 0.16, 0.02, or 0.01, with ambiguous coherence levels (0.01 or 0.02) on two thirds of trials. The coherence levels required for the direction of motion to be perceptually ambiguous were determined in a pilot study.

Following each presentation of the RDK and response, participants were provided with on-screen feedback. For the 'offered reward' trials the correct response was to continue holding the key for two seconds ('stay') to gain a monetary reward, while for the 'threatened loss' trials the correct response was to release the key prior to two seconds ('go') to avoid a monetary loss. In the first two practice blocks either 'Correct' (in green font) or 'Incorrect' (in red font) was displayed for 1500ms. In the final practice block and test block, the notional amount won (in green figures) or lost (in red figures) was displayed for 1500ms. Participants were told that the amounts they saw on the screen during these blocks would be multiplied by a factor, and then added to or deducted from an initial £2 endowment, and a £5 turn-up fee. To sustain motivation across the test session, participants were informed that the top-three ranking participants would have their bonus doubled.

In the final (third) training block, if an individual made a risky response ('stay') they would notionally win or lose £10, depending on the stimulus presented, otherwise they would get £0. For the test block, following Guitart-Masip et al. (2011), the notional amount rewarded in the fluctuating reward condition, or potential loss, in the fluctuating loss condition, for correct or incorrect stay responses respectively varied across trials according to a sine function with added noise. In the fluctuating reward condition, the reward varied between £0.87 and £19.16 and the loss was fixed at £10.

In the fluctuating loss condition, the loss varied between £0.87 and £19.16 and the reward was fixed at £10 (Fig. 3.3). The potential monetary reward and loss were displayed on screen for 1500s using connected green and red bars prior to each trial. The length of the green bar was proportional to the potential reward and the length of the red bar was proportional to the potential loss, with the amounts written in figures at the non-connecting end of each bar.

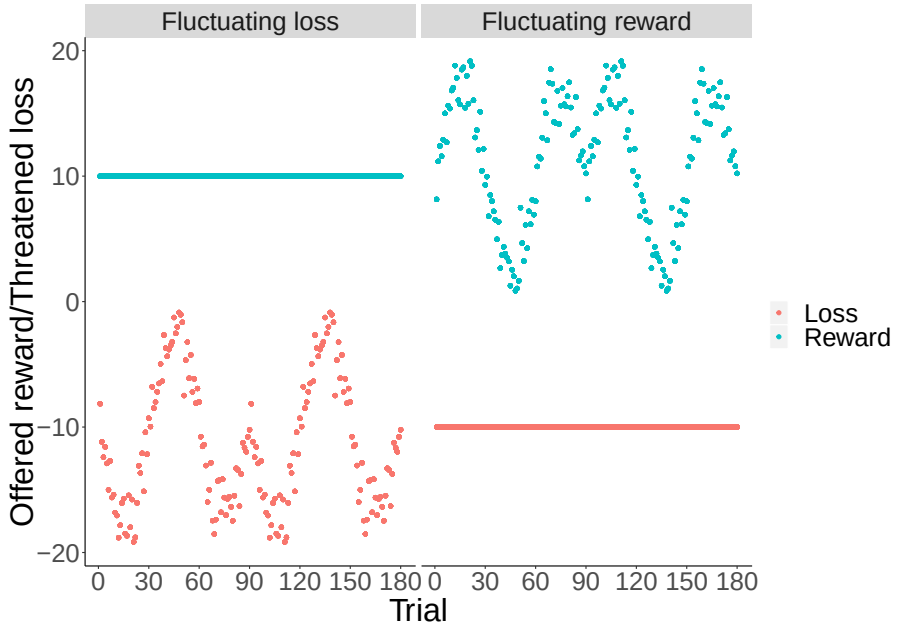
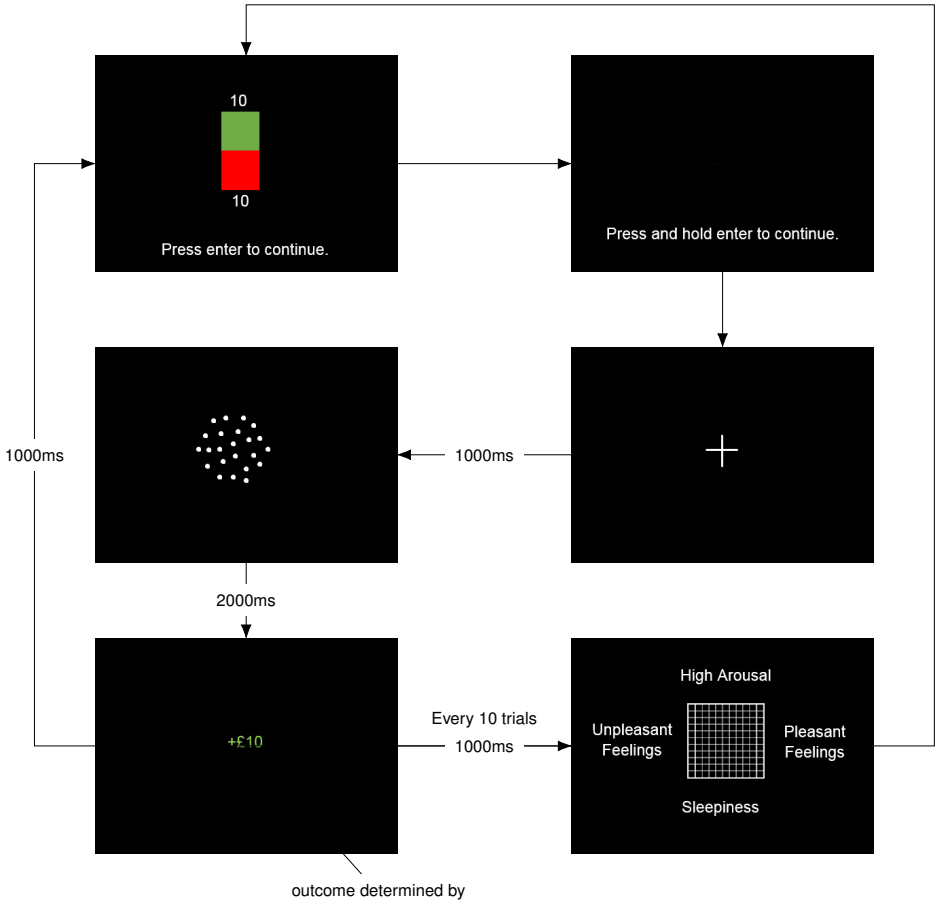


FIGURE 3.3: Offered monetary reward and threatened monetary loss on each trial of the judgement bias task in the fluctuating reward and loss conditions

3.2.4 Self reports of affect (see Fig. 3.4)

At the start of the test block of each task and following every 10 subsequent trials, participants were asked to report their current mood using a 9 by 9 computerised self-report affect grid (Killgore, 1998). To complete the affect grid participants had to move a cross, which was initially central in the grid, to the location that best described their current mood using the arrow keys on a keyboard. Horizontal movements represented changes in mood valence, with movements to the right reporting a more positively valenced mood. Vertical movements represented arousal, with upwards movement reporting higher levels of arousal.



	Stay (hold key)	Go (release key)
Favourable direction	Reward	£0
Unfavourable direction	Loss	£0

FIGURE 3.4: Structure of the human monetary judgement bias test session

3.2.5 Model dependent analysis

POMDP model

The decision-making process was modelled as a partially observable Markov decision process (POMDP) (Åström, 1965; Smallwood and Sondik, 1973) with a two-dimensional state space $s = (t, X)$ in which participants accumulate evidence (X) from observations of the RDK (x) over (veridical) time (t). The true direction of motion of the RDK (μ) takes one of two values, 1 or -1, where 1 represents motion in the favourable direction and -1 represents motion in the unfavourable direction. When $\mu = 1$, the participant should 'stay', which results in a reward, and when $\mu = -1$ the correct response is to 'go' before the trial ends, which avoids a loss. Participants have to use the evidence they collect, together with the costs and benefits of being correct

or incorrect, to decide what to do. The participants' capacity to perform interval timing of the stay period is noisy.

For convenience, we discretise the objective time between zero and two seconds into bins of Δt, and use integer states $t = \{0, 1, 2, \ldots, T\}\Delta t$. Thus, we write x_t to represent the observations from time $(t-1)\Delta t$ to $t\Delta t$, and $x_{0:t}$ to represent all the observations from the beginning of the trial up to time t. The participant's relative belief that the stimulus is favourable at time t will depend on their prior belief that the stimulus would be favourable, the relative likelihood of their observations prior to time t, and the relative likelihood of the current observation:

$$\log\left(\frac{P(\mu{=}1|x_{0:t},\theta)}{P(\mu{=}{-1}|x_{0:t},\theta)}\right) = \log\left(\frac{P(\mu{=}1)}{P(\mu{=}{-1})}\right) + \log\left(\frac{P(x_{0:t}|\mu{=}1,\theta)}{P(x_{0:t}|\mu{=}{-1},\theta)}\right) \tag{3.1}$$

$$= \log\left(\frac{P(\mu{=}1)}{P(\mu{=}{-1})}\right) + \log\left(\frac{P(x_{0:t-1}|\mu{=}1,\theta)}{P(x_{0:t-1}|\mu{=}{-1},\theta)}\right) + \log\left(\frac{P(x_t|\mu{=}1,\theta)}{P(x_t|\mu{=}{-1},\theta)}\right) \tag{3.2}$$

We assume that the likelihood follows a Gaussian probability distribution with a mean dependent on the true value of μ and the coherence level θ, and (fixed) variance σ^2 that reflects the participant's ability to detect the direction of motion of the RDK[1]: $P(x_t|\mu,\theta) \sim \mathcal{N}(x_t : \mu\theta, \sigma^2)$. Thus, the relative posterior probability for the last sample is:

$$\log\left(\frac{P(x_t|\mu{=}1,\theta)}{P(x_t|\mu{=}{-1},\theta)}\right) = \frac{2\theta}{\sigma^2}x_t \tag{3.3}$$

implying that equation the above equation can be written as:

$$\log\left(\frac{P(\mu{=}1|x_{0:t},\theta)}{P(\mu{=}{-1}|x_{0:t},\theta)}\right) = \log\left(\frac{P(\mu{=}1)}{P(\mu{=}{-1})}\right) + \frac{2\theta}{\sigma^2}\sum\nolimits_{\tau=0}^{t} x_\tau \tag{3.4}$$

Again, for convenience, we discretise the belief state X. Since the participant does not know the value of θ, we accumulate statistics $X = \sum_{\tau=0}^{t} x_\tau$ in discrete steps $\chi \times [\cdots -3, -2, -1, 0, 1, 2, 3, \ldots]$ for discretisation χ:

The state space contains two special states along with $\{t, X\}$: one, `leave`, is the inevitable consequence of choosing action GO, the other, `timeout`, arises after 2 seconds if the participant does not actively go.

We describe the probabilistic transition structure (T) of the chain in stages, i.e., the probability of going from state s to state s' when executing action a. First, we have $T_{s;\texttt{leave}}(\text{GO}) = 1$, and $T_{s;\texttt{leave}}(a) = 0, a \neq \text{GO}$. Second, we simplify the stochastic, interval-timing (Gibbon, 1977) relationship between objective and subjective time by imagining that `timeout` can happen probabilistically when the participant chooses $a =$ STAY[2]. We consider that this happens according to a gamma distribution $\Gamma(k, \phi)$, with

[1] For convenience, we suppress the dependence of μ and σ^2 on Δt.

[2] It would be more realistic for the participants' time to evolve *subjectively* rather than *objectively*, and for `timeout` to happen deterministically in objective time. However, this would mean a non-uniform acquisition of evidence (i.e. the statistics of x_t would not be homogeneous), making for extra complexities of only modest import.

shape and scale parameters k, ϕ respectively. Thus, considering the hazard function, we have:

$$P_{t;\texttt{timeout}} = P\left(\texttt{timeout} \in (t, t+1)\Delta t | \text{continuing at } t \;\&\; a_t = \textsc{stay}\right) \tag{3.5}$$

$$= \frac{\int_{t\Delta t}^{(t+1)\Delta t} d\tau\ \Gamma(\tau; k, \phi)}{1 - \int_0^{t\Delta t} d\tau\ \Gamma(\tau; k, \phi)} \tag{3.6}$$

and imposing timing out at the end of the trial by fiat:

$$P_{T;\texttt{timeout}} = 1 \tag{3.7}$$

Inclusion of the hazard function allows us to capture a core feature of the task; the participant experiences a trade-off between accruing more evidence to make a more informed decision, and ensuring that the decision is made before the end of the two-second trial (which could result in a loss). This is particularly relevant to trials where low coherence stimuli are presented as an individual may not accrue sufficient evidence to be certain about the outcome of the 'stay' response, and consequently may inadvertently make the 'stay' response by not making a decision before the end of the trial. Moreover, by considering this time pressure, the model can be easily adapted for fitting of reaction time data (see equation 3.22). The use of a hazard function based on the Gamma distribution with the the given values of ζ and ϕ is based on a well-established phenomenon in the time perception literature; the increase in variability of estimates of time as the duration of time increases (Gibbon, 1977; Matell and Meck, 2000). More specifically, the logarithmic shape of the hazard function captures the scalar properties of time perception.

Next, if the participant knows the quality of the stimulus, μ, θ, and there is no timeout when choosing STAY, we have for a single sample of x over time Δt:

$$P^{\mu,\theta}_{X;X'} = P\left(x \in \chi(X' - X) + [-\chi/2, \chi/2] | \text{no timeout}, \textsc{stay}, \mu, \theta\right) \tag{3.8}$$

$$= \left[\Phi_\sigma\left(\chi(j-i) + \frac{1}{2}\chi - \mu\theta\right) - \Phi_\sigma\left(\chi(j-i) - \frac{1}{2}\chi - \mu\theta\right)\right] \tag{3.9}$$

where Φ_σ is the cumulative normal function for $\mathcal{N}(0, \sigma^2)$.

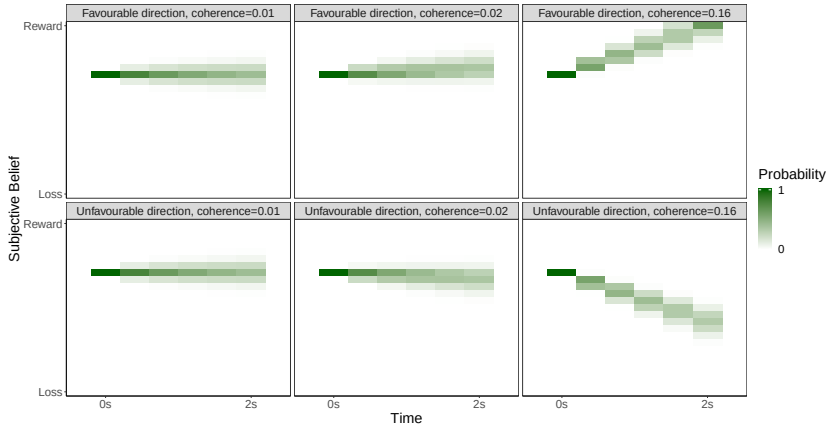
Of course, the participant does not know the true value of μ or θ. Thus, the evidence component of the subjective transition matrix $T_{s;s'}(a)$ comes from averaging over the possible μ and θ, given the information available at the current state, i.e., using the posterior probability:

$$P(\mu, \theta | t, X) = \frac{P(\mu)P(\theta)P(X|\mu, \theta, t)}{\Sigma_j \Sigma_k P(\mu)P(\theta)P(X|\mu_j, \theta_k, t)} \quad \text{, where:} \tag{3.10}$$

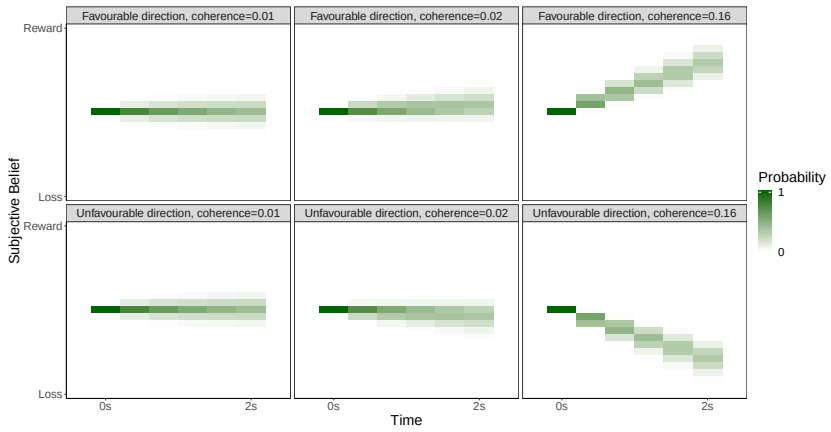
$$P(\mu = 1) = \frac{\omega}{\omega + 1} \qquad P(\mu = -1) = \frac{1}{\omega + 1} \tag{3.11}$$

$$P(\theta = 0.01) = P(\theta = 0.02) = P(\theta = 0.16) = \frac{1}{3} \tag{3.12}$$

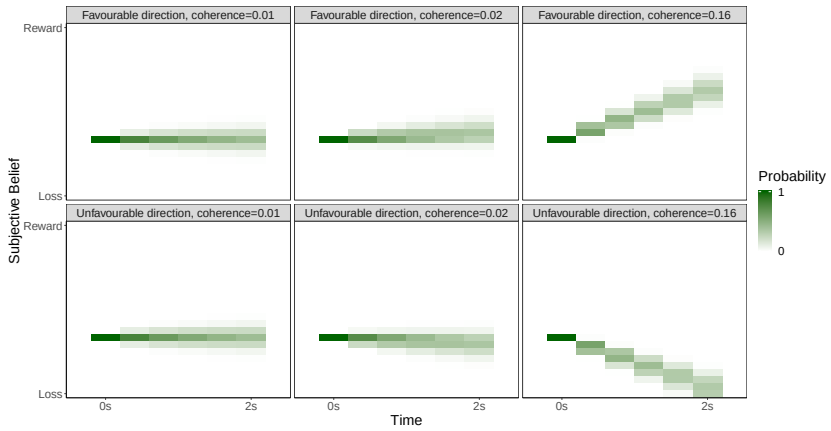
where ω is a parameter that determines the prior probability of μ, and $P(X|\mu, \theta, t)$ is determined by the discretisation of the normal distribution around $X\chi - \mu\theta t$ (Fig. 3.5).



(A) $\omega = 10$



(B) $\omega = 1$



(C) $\omega = 0.1$

FIGURE 3.5: Heatmaps illustrating the probability that a participant transitions to each state given the stimulus (i.e. true values of μ and θ) and prior (i.e. values of ω). The coherence level (θ) governs the rate of the transition towards a stronger belief that the 'stay' response will lead to a reward or loss; the direction of the stimulus (μ) governs the direction of the transitions; and the prior belief (determined by ω) governs the starting point.

In sum, the subjective transition structure from current state $s = (t, X)$ is therefore given by:

$$T_{(t,X);\texttt{leave}}(\text{GO}) = 1 \tag{3.13}$$

$$T_{(t,X);\texttt{timeout}}(\text{STAY}) = P_{t;\texttt{timeout}} \tag{3.14}$$

$$T_{(t,X);(t+1,X')}(\text{STAY}) = (1 - P_{t;\texttt{timeout}}) \left(\sum_j \sum_k P(\mu_j, \theta_k | t, X) P_{X;X'}^{\mu_j, \theta_k} \right) \tag{3.15}$$

A participant's policy π is determined by the long-run values $Q^\pi(s, a)$ of executing action a in state $s = (t, X)$, and then following the policy thereafter. In general, these action values are determined by the Bellman equation (Bellman, 1952). We assume that no discounting occurs given the short duration of the trial. In this case:

$$Q^\pi(s, \text{GO}) = 0 \tag{3.16}$$

$$Q^\pi(s, \text{STAY}) = T_{s;\texttt{timeout}}(\text{STAY}) V^{\texttt{timeout}}(s) + \sum_{X'} T_{s;s'}(\text{STAY}) V^\pi(s') \tag{3.17}$$

where $s' = (t + 1, X')$. Here, the expected timeout value depends on the likelihood of the two possibilities:

$$V^{\texttt{timeout}}(s) = C_R R_n P(\mu{=}1|s) + C_L L_n P(\mu{=}{-}1|s) \tag{3.18}$$

where C_R and C_L are parameters that reflect the participant's sensitivity to gains and losses respectively, and R_n and L_n are the potential reward and loss on trial n. Further:

$$V^\pi(s') = \pi(s', \text{GO}') \times 0 + \pi(s', \text{STAY}) Q^\pi(s', a) \tag{3.19}$$

where π is determined by:

$$\pi(s', \text{STAY}) = \lambda + (1 - 2\lambda)\sigma(B[Q^\pi(s', \text{STAY}) - Q^\pi(s', \text{GO})]); \qquad \pi(s', \text{GO}') = 1 - \pi(s', \text{STAY}) \tag{3.20}$$

with $\sigma(z) = 1/(1 + \exp(-z)$ being the logistic sigmoid, and λ and B being lapse and inverse temperature parameters respectively (Fig. 3.6).

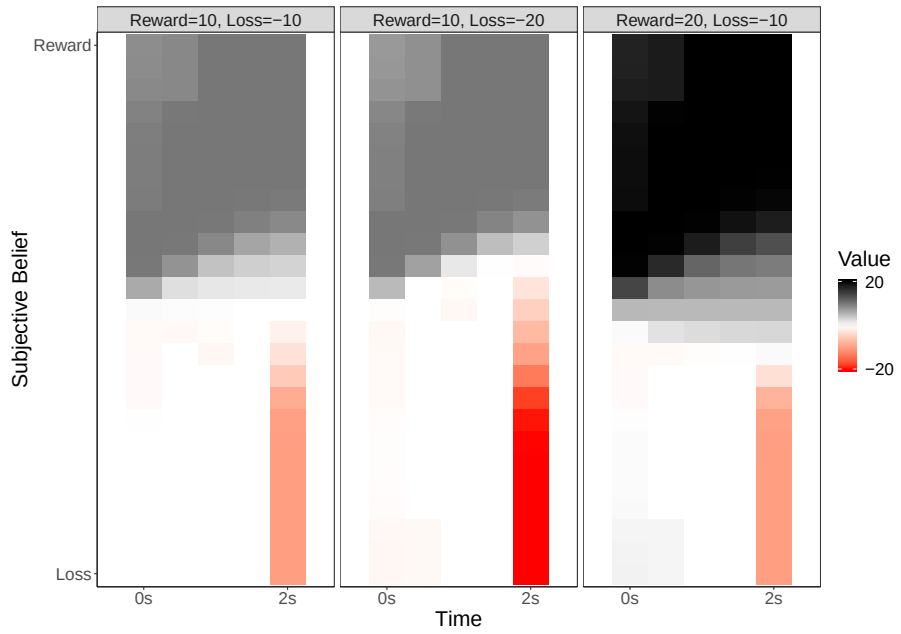


FIGURE 3.6: Heatmap illustrating the value of occupying each state as the value of the subjective reward and loss varies; this could either result from variation in the true value of the reward and loss, or fluctuations in C_R or C_L which scale the true value of the reward and loss. The state values earlier in the trial in which there is a weaker belief that the trial will be rewarded (i.e. bottom y-axis; left-side x-axis) are close to zero; this reflects that the 'leave' action is the most likely future action and hence the trial outcome will most likely be zero. The value of states representing no strong belief about the outcome of the trial (i.e. central y-axis) changes across time; this reflects that transitions to a state with a high value (e.g. resulting from a strong certainty of reward) are possible earlier in the trial, but if there is uncertainty about the stimulus further into the trial, then it is most likely that the stimulus has a low coherence and that reaching a high value state is unlikely.

To calculate the distribution of leaving times, we consider the probability $P_s^{\texttt{remain}}(\mu_n, \theta_n)$ of staying until state $s = (t, X)$ on trial n, which is defined by direction μ_n and coherence θ_n. This enjoys a recursive form:

$$P_{0,X}^{\texttt{remain}}(\mu_n, \theta_n) = 1 \tag{3.21}$$

$$P_{t,X}^{\texttt{remain}}(\mu_n, \theta_n) = \int_{x_t} dx_t \; P_{t-1,X-x_t}^{\texttt{remain}}(\mu_n, \theta_n)\pi(\text{STAY}, (t-1, X-x_t))(1 - P_{t-1;\texttt{timeout}})P(x_t|\mu_n, \theta_n) \tag{3.22}$$

from which we can calculate the overall probability of remaining until timing out as:

$$P^{\texttt{timeout}}(\mu_n, \theta_n) = \int_X dX \; \sum_{t=0}^{t} = 2sP_{t,X}^{\texttt{remain}}(\mu_n, \theta_n)\pi(\text{STAY}, (t, X))P_{t;\texttt{timeout}} \tag{3.23}$$

In this framework, individual differences in decision-making, which will appear as biases, arise through variation in C_R, C_L, and ω. As we hypothesise that judgement

bias will depend on past experience, specifically the average earning rate ($\bar{R}_n$), the weighted (low pass filtered) prediction error ($w\text{PE}_n$), the weighted (low pass filtered) squared prediction error ($w\text{PE}_n^2$), and the most recent outcome (O_{n-1}) we allow C_R, C_L, and ω potentially to depend additively on these values as well as on a constant term, which reflects baseline individual variation in these parameters. This is mapped through an exponential so that the values of C_R, C_L, and ω are always positive. To allow the relative contribution of $\bar{R}_{n-1}$, $w\text{PE}_{n-1}$, $w\text{PE}_{n-1}^2$, O_{n-1} to C_R, C_L, and ω to vary, we scale $\bar{R}_{n-1}$, $w\text{PE}_{n-1}$, $w\text{PE}_{n-1}^2$, and O_{n-1} by weighting parameters:

$$\omega = e^{\left(\beta_0^{\omega}+\beta_{\bar{R}}^{\omega}\bar{R}_{n-1}+\beta_{w\text{PE}}^{\omega}w\text{PE}_{n-1}+\beta_{w\text{PE}^2}^{\omega}w\text{PE}_{n-1}^2+\beta_O^{\omega}O_{n-1}\right)} \tag{3.24}$$

$$C_R = e^{\left(\beta_0^{C_R}+\beta_{\bar{R}}^{C_R}\bar{R}_{n-1}+\beta_{w\text{PE}}^{C_R}w\text{PE}_{n-1}+\beta_{w\text{PE}^2}^{C_R}w\text{PE}_{n-1}^2+\beta_O^{C_R}O_{n-1}\right)} \tag{3.25}$$

$$C_L = e^{\left(\beta_0^{C_L}+\beta_{\bar{R}}^{C_L}\bar{R}_{n-1}+\beta_{w\text{PE}}^{C_L}w\text{PE}_{n-1}+\beta_{w\text{PE}_n^2}^{C_L}w\text{PE}_{n-1}^2+\beta_O^{C_L}O_{n-1}\right)} \tag{3.26}$$

$\bar{R}_{n-1}$ and $w\text{PE}_{n-1}$ are updated following the outcome of each trial O_n (which is scaled by reward or loss sensitivity such that O_n is $C_R R_n$, $C_L L_n$, or zero). Specifically, following Rutledge et al. (2014), $w\text{PE}_{n-1}$ is the prediction error on past trials $\text{PE}_{1:n-1}$ weighted such that the influence of past prediction errors attenuates over trials:

$$w\text{PE}_n = \frac{1-\gamma_{w\text{PE}}}{1-\gamma_{w\text{PE}}^n}\Sigma_{i=1}^{n}\gamma_{w\text{PE}}^{n-i}\text{PE}_i \tag{3.27}$$

Participants may make predictions about the outcome of a trial both prior to and following stimulus presentation, but it is unclear which of these predictions would lead to prediction errors most relevant to future decision-making. We assess these two possibilities by fitting a set of models in which PE_n is defined as the difference between O_n and expected outcome of the trial prior to stimulus presentation, as given by the average of the value function:

$$\text{PE}_n = O_n - V(0,0) \tag{3.28}$$

Then, comparing this set of models to a set of models in which PE_n is zero where the 'go' response is made, and the difference between O_n and the stimulus-dependent expected value of the state when the trial terminates:

$$\text{PE}_n = \begin{cases} 0, & \text{if action = leave} \\ O_n - \left[\sum_s^{t=2s} P(t|X=\texttt{timeout})V(s)P(x_t|\mu_n,\theta_n)\right], & \text{otherwise} \end{cases} \tag{3.29}$$

Where $P(t|X=\texttt{timeout})$ is:

$$P(t|X = \texttt{timeout}) = \frac{P(X = \texttt{timeout}|t)}{\Sigma_{j=1} P(X = \texttt{timeout}|t_j)} \tag{3.30}$$

Overall, a prediction error based on a prediction prior to the presentation of the stimulus was found to provide the best model fit (ΔAIC=301.457) The average earning rate ($\alpha_{\bar{R}_n}$), which reflects the learnt value of the test session from previous wins and losses, updates according to a Rescorla-Wagner learning model, with learning rate $\alpha_{\bar{R}_n}$ (Rescorla et al., 1972):

$$\bar{R}_n = \bar{R}_{n-1} + \alpha_{\bar{R}}(O_n - \bar{R}_{n-1}) \tag{3.31}$$

The key parameters in the model were thus (Fig. 3.7 and Table 3.1): σ, λ, β_0^{ω}, $\beta_0^{C_R}$, $\beta_0^{C_L}$, $\beta_{\bar{R}}^{\omega}$, $\beta_{\bar{R}}^{C_R}$, $\beta_{\bar{R}}^{C_L}$, β_{wPE}^{ω}, $\beta_{wPE}^{C_R}$, $\beta_{wPE}^{C_L}$, $\beta_{wPE^2}^{\omega}$, $\beta_{wPE^2}^{C_R}$, $\beta_{wPE^2}^{C_L}$, β_O^{ω}, $\beta_O^{C_R}$, $\beta_O^{C_L}$, γ_{wPE} and $\alpha_{\bar{R}}$, although these were not fitted simultaneously (see below section). The remaining parameters which were of no interest here (i.e. B, k, and ϕ) were fixed, with their values based on sensible estimates (B=20.086, k=5, and ϕ=1).

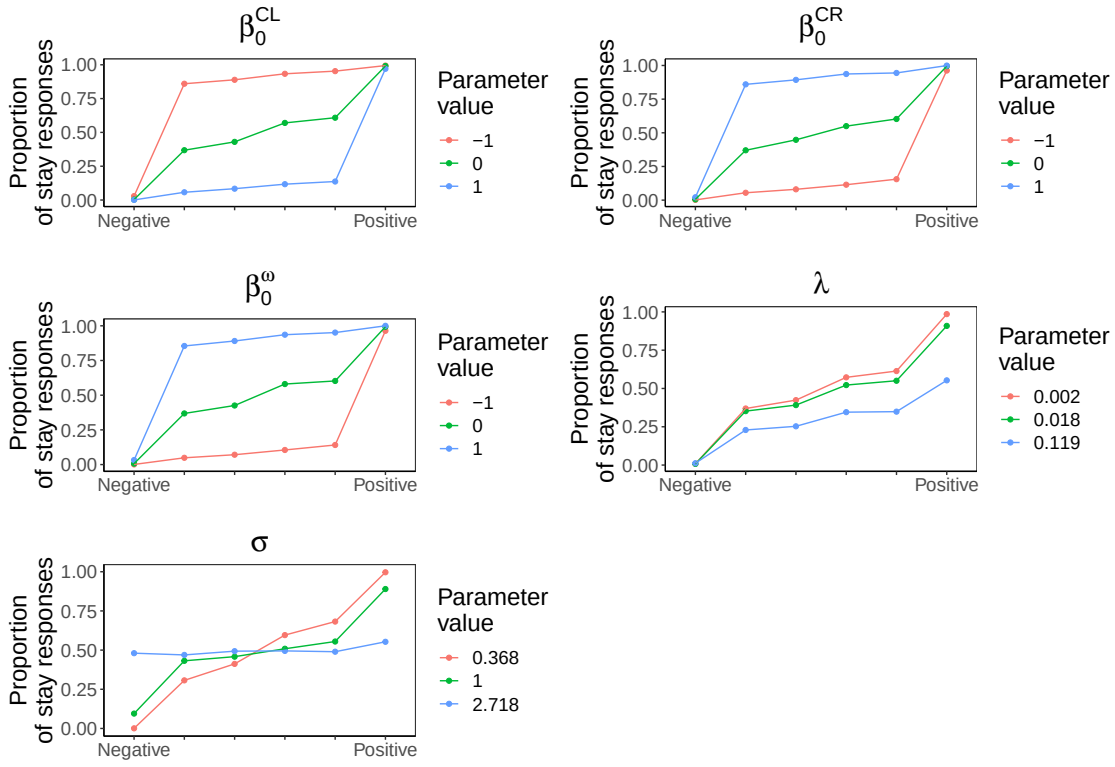


FIGURE 3.7: Plots of judgement bias data generated using the described model in which the values of $\beta_0^{C_L}$, $\beta_0^{C_R}$, β_0^{ω}, λ and σ are varied. $\beta_0^{C_L}$, $\beta_0^{C_R}$, β_0^{ω} each govern the extent to which participants make the 'stay' or 'go' response, with $\beta_0^{C_L}$ exerting an opposite effect to $\beta_0^{C_R}$ and β_0^{ω}; λ governs the accuracy of response – the asymmetry results from there being several opportunities to make a 'go' response after a 'stay' response, but no further opportunities to make the 'stay' response after a 'go' response; σ governs the extent to which participants can discriminate between the stimuli, and hence accuracy.

Parameter	Range	Interpretation
$\alpha_{\bar{R}}$	$[0, 1]$	Learning rate for the average earning rate; higher values reflect faster updating of the average earning rate.
B	$[0, \infty]$	Inverse temperature parameter; higher values reflect that decisions are less stochastic and more heavily based on the expected value of actions.
$\beta_0^{C_L}$	$[-\infty, \infty]$	Loss sensitivity baseline; higher values reflect greater that losses are overall weighted more heavily.
$\beta_{\bar{R}}^{C_L}$	$[-\infty, \infty]$	Average earning rate dependent loss sensitivity; higher values reflect that losses are weighted more heavily when the average earning rate is higher.
$\beta_O^{C_L}$	$[-\infty, \infty]$	Outcome dependent loss sensitivity; higher values reflect that losses are weighted more heavily when the most recent outcome is higher.
$\beta_{w\mathrm{PE}}^{C_L}$	$[-\infty, \infty]$	Weighted prediction error dependent loss sensitivity; higher values reflect that losses are weighted more heavily when the weighted prediction error is higher.
$\beta_{w\mathrm{PE}^2}^{C_L}$	$[-\infty, \infty]$	Weighted squared prediction error dependent loss sensitivity; higher values reflect that losses are weighted more heavily when the weighted squared prediction error is higher.
$\beta_0^{C_R}$	$[-\infty, \infty]$	Reward sensitivity baseline; higher values reflect greater that rewards are overall weighted more heavily.
$\beta_{\bar{R}}^{C_R}$	$[-\infty, \infty]$	Average earning rate dependent reward sensitivity; higher values reflect that rewards are weighted more heavily when the average earning rate is higher.
$\beta_O^{C_R}$	$[-\infty, \infty]$	Outcome dependent reward sensitivity; higher values reflect that rewards are weighted more heavily when the most recent outcome is higher.
$\beta_{w\mathrm{PE}}^{C_R}$	$[-\infty, \infty]$	Weighted prediction error dependent reward sensitivity; higher values reflect that rewards are weighted more heavily when the weighted prediction error is higher.
$\beta_{w\mathrm{PE}^2}^{C_R}$	$[-\infty, \infty]$	Weighted squared prediction error dependent reward sensitivity; higher values reflect that rewards are weighted more heavily when the weighted squared prediction error is higher.
β_n	$[-\infty, \infty]$	Time-dependency of the policy: higher values reflect that the 'stay' action is less likely as more trials are completed.
β_0^{ω}	$[-\infty, \infty]$	Prior baseline; higher values reflect greater that rewards are overall considered more likely that losses.
$\beta_{\bar{R}}^{\omega}$	$[-\infty, \infty]$	Average earning rate dependent prior; higher values reflect a greater prior belief about reward when the average earning rate is higher.
β_O^{ω}	$[-\infty, \infty]$	Outcome dependent prior; higher values reflect a greater prior belief about reward when the previous outcome is higher.
$\beta_{w\mathrm{PE}}^{\omega}$	$[-\infty, \infty]$	Weighted prediction error dependent prior; higher values reflect a greater prior belief about reward when the weighted prediction error is higher.
$\beta_{w\mathrm{PE}^2}^{\omega}$	$[-\infty, \infty]$	Weighted squared prediction error dependent prior; higher values reflect a greater prior belief about reward when the weighted squared prediction error is higher.
$\gamma_{w\mathrm{PE}}$	$[0, 1]$	Forgetting factor for the weighted prediction error; higher values reflect slower forgetting of previous prediction errors.
ζ	$[0, \infty]$	Shape parameter for the Gamma function determining the hazard function for the likelihood of timeout; values lower than one indicate that timeout is more likely at earlier timesteps, while the steepness of the function increases from values greater than one (i.e. timeout becomes increasingly likely with increasing timesteps).
λ	$[0, 1]$	Lapse rate: higher values a greater likelihood of executing the action with the lowest value (i.e. the 'wrong' action).
σ	$[0, \infty]$	Slope parameter: higher values reflect a poorer ability to detect the true direction of the stimulus.
ϕ	$[0, \infty]$	Scale parameter for the Gamma function determining the hazard function for the likelihood of timeout; higher values decrease the likelihood of timeout across all timesteps.

Table 3.1: Glossary of model parameters

However, due to significant correlations between $w\mathrm{PE}_{n-1}$ $w\mathrm{PE}^2_{n-1}$, and O_{n-1}, with the extent of the correlation dependent on the model parameters, parameters characterising the influence of these variables on the same aspect of the decision-making process (e.g. $\beta^{\omega}_{w\mathrm{PE}}$, $\beta^{\omega}_{w\mathrm{PE}^2}$, and β^{ω}_{O}: $\beta^{C_R}_{w\mathrm{PE}}$, $\beta^{C_R}_{w\mathrm{PE}^2}$, and $\beta^{C_R}_{O}$; $\beta^{C_L}_{w\mathrm{PE}}$, $\beta^{C_L}_{w\mathrm{PE}^2}$, and $\beta^{C_L}_{O}$) were not fitted simultaneously in the same model, but instead fitted separately and the goodness of fit of each model compared using their AIC values.

Model-fitting and analysis of parameter estimates

We fitted the data for a given model to each subject's data individually by maximum likelihood, adding parameters in stepwise manner, and compared models using Akaike's information criterion (AIC) values. Thus, parameters that did not increase the parsimony of the model were excluded. The model was found to provide accurate recovery of parameters, as determined by simulating data and assessing the correlations between the parameters recovered by the model and those used to simulate the data (see Appendix B).

Model-fitting was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bris.ac.uk/acrc/. This involved using 16 nodes of the high-performance computer (BlueCrystal Phase 3) to fit each model to each participant's data in parallel, allowing each model to be fitted to the whole dataset in under 24 hours. The likelihood function and associated code for the model was written using MATLAB and the model was fitted to the data using the MATLAB function fmincon. The initial values are shown in Table 3.2. Values for parameters constrained between zero and ∞ were exponentiated within the likelihood function, and likewise values for parameters constrained between zero and one were transformed using a logistic function within the likelihood function.

TABLE 3.2: Initial values for the model fitting procedure using fmincon. Note that these values are transformed within the likelihood function.

Parameter	**Initial value**
$\alpha_{\bar{R}}$	-4
B	3
$\beta_0^{C_L}$	0
$\beta_{\bar{R}}^{C_L}$	0
$\beta_O^{C_L}$	0
$\beta_{w\mathrm{PE}}^{C_L}$	0
$\beta_{w\mathrm{PE}^2}^{C_L}$	0
$\beta_0^{C_R}$	0
$\beta_{\bar{R}}^{C_R}$	0
$\beta_O^{C_R}$	0
$\beta_{w\mathrm{PE}}^{C_R}$	0
$\beta_{w\mathrm{PE}^2}^{C_R}$	0
β_n	0
β_0^{ω}	0
$\beta_{\bar{R}}^{\omega}$	0
β_O^{ω}	0
$\beta_{w\mathrm{PE}}^{\omega}$	0
$\beta_{w\mathrm{PE}^2}^{\omega}$	0
$\gamma_{w\mathrm{PE}}$	-4
ζ	1.609
λ	-10
σ	-0.7
ϕ	0

The aim of the model-dependent analysis was to investigate the relationship between reward and loss experience and decision-making, specifically to examine which, if any, cognitive processes involved in decision-making were modulated by different aspects of reward and loss experience. Hence, analysis of the parameter estimates from the POMDP model examined the key predictions of this study.

We used permutation tests (PT) to assess whether the parameter estimates from the most parsimonious model as determined in model-fitting differed significantly from zero. To examine whether the parameter estimates correlated with reported affect or condition, and hence to examine how affect might mediate the relationship between reward and punisher experience and decision-making, a general linear model was fitted to the parameter estimates with mean reported arousal (from the affect grid data), mean reported valence (from the affect grid data), and condition as the predictor variables. Visual inspection of the model residuals verified that the assumptions of each model were met.

3.2.6 Model-agnostic statistical analysis

Generalised linear mixed models (GLMMs) were fitted to the affect grid, trial initiation, and judgement bias data in R (R Core Team, 2017) using the lme4 (Bates et al., 2015) and nlme (Pinheiro et al., 2018) packages and likelihood ratio tests were then used to assess whether the difference in model deviance was significant following removal of a parameter from a model. Reported affective valence and arousal as well as trial initiation data were log-transformed to ensure that assumptions of normality and homogeneity of variance were not violated. Visual inspection of the model residuals verified that the assumptions of each model were met.

Analysis of the affect grid data allowed us to examine the relationship between reward and loss experience and affect; a key aim of this chapter. Analysis of the trial initiation data allowed us to examine whether reward and loss experience altered vigour; a secondary aim of this chapter. These model agnostic analyses also allowed us to assess whether there were overall differences in affect, vigour, or decision-making between the fluctuating reward and loss conditions; whether there were time-dependent changes in affect, vigour, and decision-making; and assess the extent to which participants were attending to the offered reward and threatened loss.

Each GLMM included a random effect of participant and fixed effects of number of trials completed, the potential outcome ($\frac{R_n+L_n}{2}$), and condition. In addition to the variables described, the model of judgement bias included a fixed effect of the RDK presented (the combined coherence and direction). To examine if and how affect and trial initiation latency depended on reward and punisher experience, $\bar{R}_{n-1}$, $w\text{PE}_{n-1}$, $w\text{PE}^2_{n-1}$, and O_{n-1} were extracted from the model which best explained the judgement bias data, and included as predictor variables in the GLMMs of initiation latency, reported valence, and reported arousal. Due to significant correlations between $w\text{PE}_{n-1}$, $w\text{PE}^2_{n-1}$, and O_{n-1}, these variables were not included in the same GLMM but instead separate GLMMs including each variable were compared according to their AIC value and the GLMM which provided the best fit was selected for further analysis. In the GLMM of trial initiation the best-fitting predictor variable was $w\text{PE}_{n-1}$ (ΔAIC=6.410, comparing saturated model containing $w\text{PE}_{n-1}$ and not $w\text{PE}^2_{n-1}$ or O_{n-1} with the next best-fitting saturated model which contained O_{n-1} but not $w\text{PE}_{n-1}$ or $w\text{PE}^2_{n-1}$), in the model of reported arousal it was $w\text{PE}^2_{n-1}$ (ΔAIC=0.260, ΔAIC=6.410, comparing saturated model containing $w\text{PE}^2_{n-1}$ and not $w\text{PE}_{n-1}$ or O_{n-1} with the next best-fitting saturated model which contained $w\text{PE}_{n-1}$ but not O_{n-1} or$w\text{PE}^2_{n-1}$), and in the model of reported valence this was O_{n-1} (ΔAIC=7.238; comparing saturated model containing O_{n-1} and not $w\text{PE}_{n-1}$ or $w\text{PE}^2_{n-1}$ with the next best-fitting saturated model which contained $w\text{PE}_{n-1}$ but not O_{n-1} or $w\text{PE}^2_{n-1}$).

The GLMMs predicting affect and trial initiation also included interaction terms between each of the variables encompassing reward and punisher experience and condition. To investigate significant and marginally non-significant interaction terms, the data were split by condition and further GLMMs were fitted to these subsetted data.

The p-values obtained from these additional GLMMs were adjusted using the false discovery rate method to account for multiple comparisons.

3.3 Results

3.3.1 Model-dependent results

Judgement bias: choice

Variation in judgement bias might arise through an altered sensitivity to rewards or losses or a prior belief that the reward was more or less likely. Individually, inclusion of parameters characterising reward sensitivity ($\beta_0^{C_R}$, ΔAIC=233.834; comparing model fitting $\beta_0^{C_R}$, λ, and σ with model fitting λ and σ), loss sensitivity ($\beta_0^{C_L}$, ΔAIC=275.350; comparing model fitting $\beta_0^{C_L}$, λ, and σ with model fitting λ and σ), and prior belief (β_0^{ω}, ΔAIC=313.037; comparing model fitting β_0^{ω}, λ, and σ with model fitting λ and σ) was found to improve the model fit (resulted in decreased AIC values) confirming that variation in these parameters could explain variation in judgement bias.

To examine whether and how different aspects of reward and punisher experience, specifically the average reward rate $\bar{R}_{n-1}$, prediction error $w\text{PE}_{n-1}$, and squared prediction error $w\text{PE}^2_{n-1}$, and the previous outcome O_{n-1}, influenced judgement bias we allowed reward sensitivity, loss sensitivity, and prior belief to depend on these values. The model fit did not improve when the prior belief depended on $\bar{R}_{n-1}$, $w\text{PE}^2_{n-1}$, $w\text{PE}_{n-1}$, or O_{n-1} (individual inclusion of $\beta^{\omega}_{\bar{R}}$, $\beta^{\omega}_{w\text{PE}^2}$, $\beta^{\omega}_{w\text{PE}}$, and β^{ω}_{O}), it improved when reward sensitivity depended on $w\text{PE}^2_{n-1}$ or $w\text{PE}_{n-1}$, (individual inclusion of $\beta^{C_R}_{w\text{PE}^2}$ and $\beta^{C_R}_{w\text{PE}}$) but not $\bar{R}_{n-1}$ or O_{n-1} (individual inclusion of $\beta^{C_R}_{\bar{R}}$ and $\beta^{C_R}_{O}$), and it also improved when loss sensitivity depended on $\bar{R}_{n-1}$ or $w\text{PE}^2_{n-1}$ (individual inclusion of $\beta^{CL}_{\bar{R}}$, $\beta^{CL}_{w\text{PE}^2}$), but not $w\text{PE}_{n-1}$ or O_{n-1} (individual inclusion of $\beta^{CL}_{w\text{PE}}$ and β^{CL}_{O}), Table 3.3.

Inclusion of a variable learning rate for $\bar{R}_{n-1}$ and variable forgetting factor for $w\text{PE}_{n-1}$ was not found to be necessary. Instead, a fixed learning rate for $\bar{R}_{n-1}$ ($\alpha_{\bar{R}}$=0.018, ΔAIC=79.648; comparing all models in which decision-making depended on $\bar{R}$ and $\alpha_{\bar{R}}$ was allowed to vary, to the same set of models in which $\alpha_{\bar{R}}$ was fixed) which allowed slow updating of the average earning rate based on that reported by Guitart-Masip et al. (2011), and fixed forgetting factor for $w\text{PE}_{n-1}$ ($\gamma_{w\text{PE}}$=0.018, ΔAIC=768.885; comparing all models in which decision-making depended on wPE and $\gamma_{w\text{PE}}$ was allowed to vary, to the same set of models in which $\gamma_{w\text{PE}}$ was fixed) based on that reported by Rutledge et al. (2014) which allowed rapid forgetting of recent prediction errors resulted in more parsimonious models.

Further models were then fitted to the data to assess whether the experience-dependent parameters characterising biases jointly contributed to variation in judgement bias (Table 3.3). The model with the lowest AIC included the following parameters: σ, λ, $\beta_0^{C_L}$, and $\beta^{C_L}_{w\text{PE}^2}$. The fit of this model is demonstrated in Fig. 3.8. The

model-derived probability of making a 'stay' response on each trial was found to be a strongly significant predictor of the observed response when analysed using a binomial GLMM with a random effect of subject (Fig. 3.9: LRT=3285, p<0.001). Thus, judgement bias was modulated by the squared weighted prediction error through variation in loss sensitivity.

TABLE 3.3: ΔAIC values for computational models of judgement bias choice data: the difference between the AIC values of each model and the AIC value for the best model

Model parameters	**ΔAIC**
σ, λ, $\beta_0^{C_L}$, $\beta_{w\mathrm{PE}^2}^{C_L}$	0.000
σ, λ, β_0^{ω}, $\beta_0^{C_L}$, $\beta_{w\mathrm{PE}^2}^{C_L}$	45.879
σ, λ, β_0^{ω}	92.501
σ, λ, β_0^{ω}, β_O^{ω}	95.778
σ, λ, β_0^{ω},$\beta_{w\mathrm{PE}}^{\omega}$	98.442
σ, λ, $\beta_0^{C_L}$, $\beta_{\bar{R}}^{C_L}$, $\beta_{w\mathrm{PE}^2}^{C_L}$	102.535
σ, λ, $\beta_0^{C_L}$, $\beta_{\bar{R}}^{C_L}$	105.404
σ, λ, $\beta_0^{C_R}$, $\beta_{w\mathrm{PE}^2}^{C_R}$	122.224
σ, λ, β_0^{ω}, $\beta_{\bar{R}}^{\omega}$	122.853
σ, λ, $\beta_0^{C_L}$	130.189
σ, λ, $\beta_0^{C_L}$,$\beta_{w\mathrm{PE}}^{C_L}$	130.773
σ, λ, $\beta_0^{C_R}$, $\beta_{w\mathrm{PE}}^{C_R}$	145.073
σ, λ, $\beta_0^{C_R}$	171.704
σ, λ, $\beta_0^{C_R}$, $\beta_{\bar{R}}^{C_R}$	220.694
σ, λ, β_0^{ω}, $\beta_{w\mathrm{PE}^2}^{\omega}$	224.933
σ, λ, $\beta_0^{C_L}$, $\beta_O^{C_L}$	250.035
σ, λ, $\beta_0^{C_L}$, $\beta_{w\mathrm{PE}^2}^{C_L}$, $\beta_0^{C_R}$, $\beta_{w\mathrm{PE}^2}^{C_R}$	294.042
σ, λ, $\beta_0^{C_R}$, $\beta_O^{C_R}$	304.957
σ, λ, β_0^{ω}, $\beta_0^{C_R}$, $\beta_{w\mathrm{PE}^2}^{C_R}$	320.442
σ, λ	405.538
λ	1842.314
σ	2598.177

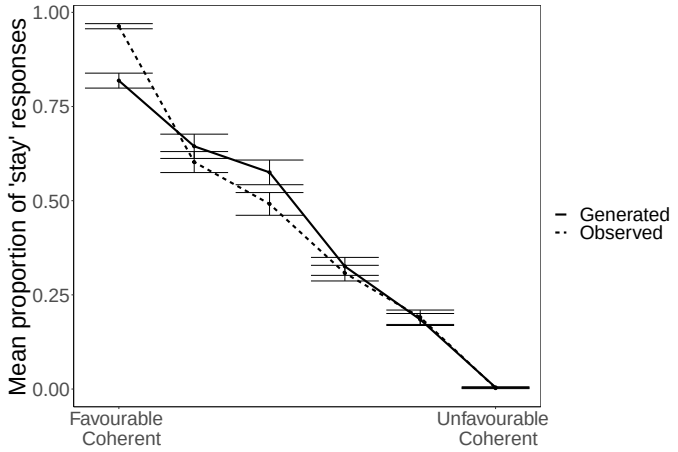


FIGURE 3.8: The mean proportion of 'stay' responses for each stimulus level for both the model-generated and observed judgement bias data. Error bars represent one standard error

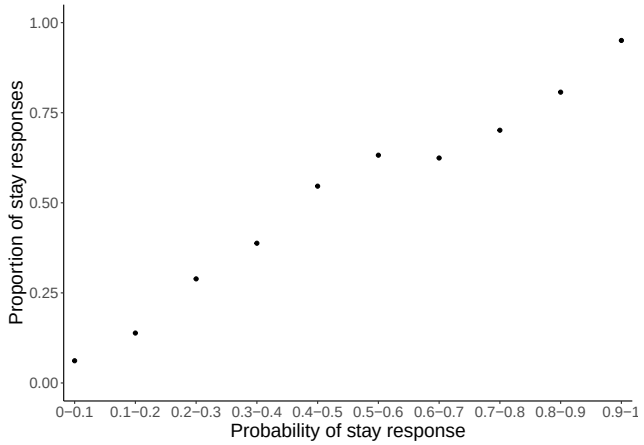


FIGURE 3.9: The proportion of 'stay' responses for intervals of model-derived probabilities of executing the 'stay' response

The estimates of $\beta_0^{C_L}$ (PT; mean±SE=-0.299±0.095, p=0.003) were significantly lower than zero, indicating that participants had a reduced value of the loss relative to the true value. Participants were also more sensitive to losses when $w\text{PE}^2$ was higher ($\beta_{w\text{PE}^2}^{C_L}$, PT: mean±SE=0.004±0.001, p=0.001).

The analysis revealed that these effects were associated with self-reported affect. Specifically, more negative values of $\beta_0^{C_L}$ (representing weaker overall loss sensitivity) were significantly associated with greater reported arousal (LRT=3.828, p=0.050), but not reported valence (LRT=0.594, p=0.441), whereas, estimates of $\beta_{w\text{PE}^2}^{C_L}$ (representing the extent to which the squared prediction error modulated loss sensitivity) were significantly correlated with reported valence (LRT=4.623, p=0.032) but not arousal (LRT=1.401, p=0.236). Participants that reported more positive affective valence overall had a loss sensitivity that was modulated to a greater extent by $w\text{PE}_{n-1}^2$, with greater loss aversion when $w\text{PE}_{n-1}^2$ was higher. The estimates of $\beta_{w\text{PE}^2}^{C_L}$ were significantly greater in the reward condition (LRT=3.845, p=0.050) and the estimates of

$\beta_0^{C_L}$ did not depend on condition (LRT=0.075, p=0.785).

3.3.2 Model-agnostic statistical analysis results

Judgement Bias

As anticipated, both the stimulus presented (LRT=3520.530, p<0.001) and the potential outcome (Fig. 3.10: LRT=5.361, p=0.021) on each trial were significant predictors of judgement bias. More specifically, participants were significantly more likely to make the 'go' response when the RDK presented was moving more coherently in the unfavourable direction and when the potential outcome (the averaged offered reward and threatened loss) was lower. Participants became more risk-averse as the test session progressed (LRT=15.727, p<0.001). Judgement bias did not differ significantly between the fluctuating reward and loss condition (LRT=1.978, p=0.160).

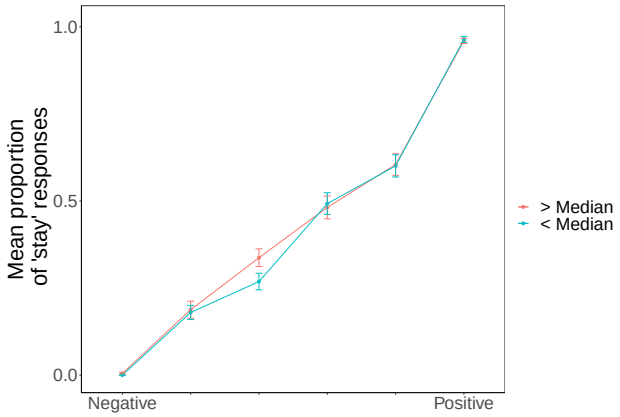


FIGURE 3.10: The mean proportion of 'stay' responses for each stimulus split by whether the potential outcome of the trial was higher (red) or lower (blue) than the median value.

Trial initiation latency

While trial initiation latencies were significantly shorter when $\bar{R}_{n-1}$ (the average earning rate) was higher (LRT=12.503, p<0.001), trial initiation latencies were significantly longer when $w\text{PE}_{n-1}$ (weighted prediction error) was more positive (LRT=65.976, p<0.001). The number of trials completed (LRT=291.277, p<0.001) was also a significant predictor of trial initiation latency; individuals were faster to initiate trials as the number of trials completed increased. Although the potential outcome (LRT=1.572, p=0.210) was not significant as a main effect and there was no overall difference in trial initiation latency between the fluctuating reward and fluctuating loss condition (LRT=0.255, p=0.614), there was a significant interaction between the potential outcome and condition (Fig. 3.11: LRT=11.690, p=0.001) as well as between $\bar{R}_{n-1}$ and condition (Fig. 3.11: LRT=41.299, p=0.000), and $w\text{PE}_{n-1}$ and condition (Fig. 3.11: LRT=11.169, p=0.001).

While participants initiated trials more slowly when the potential outcome was higher (potential reward was more positive) in the fluctuating reward condition (LRT=8.185, p=0.006), they tended to initiate trials more slowly when the potential outcome was lower (potential loss was more negative) in the fluctuating loss condition (LRT=2.730, 0.098). Although participants were faster to initiate trials when $\bar{R}_{n-1}$ and $w\text{PE}_{n-1}$ were more positive in both the fluctuating reward ($\bar{R}_{n-1}$: LRT=9.656, p=0.004; $w\text{PE}_{n-1}$: LRT=55.305, p<0.001) and loss condition ($\bar{R}_{n-1}$: LRT=3.546, p=0.072; $w\text{PE}_{n-1}$: LRT=15.801,p<0.001), the regression weight associated with the GLMM indicated that these effects were greater in the fluctuating reward condition.

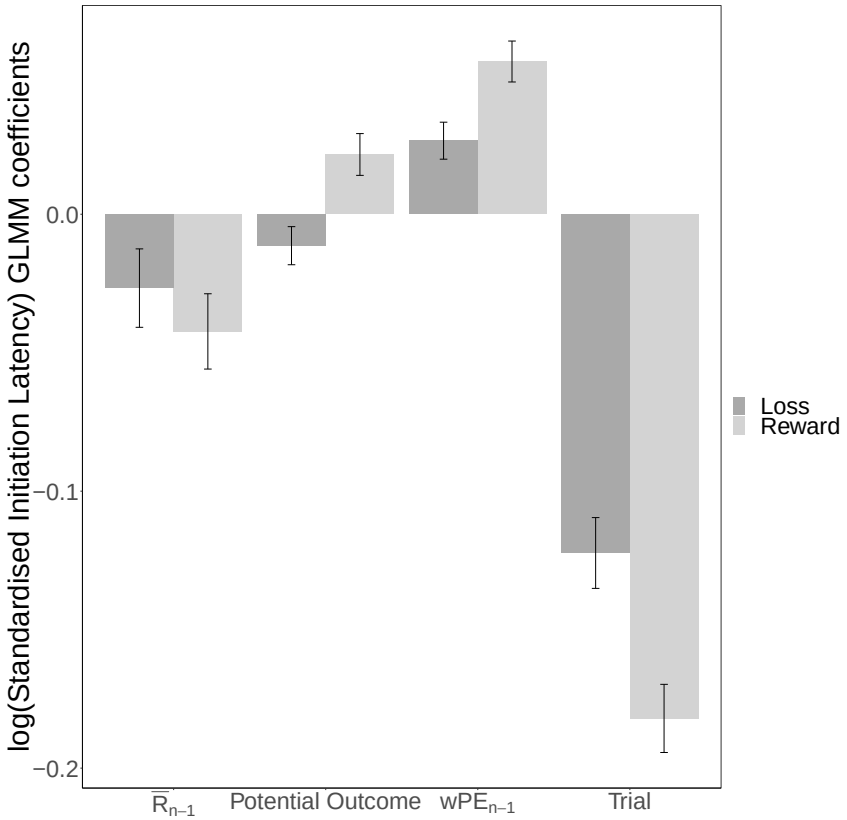


FIGURE 3.11: Standardised GLMM coefficients from the model of the log-transformed latency to initiate each trial for both the fluctuating reward and fluctuating loss condition. Error bars represent one standard error

Reported Affect

The analysis of reported affective valence showed that several aspects of reward and punisher experience influenced reported affective valence. Participants reported significantly more positive affective valence when $\bar{R}_{n-1}$ (average earning rate; LRT=42.767, p=<0.001), O_{n-1} (previous outcome; LRT=12.460, p<0.001), and the potential outcome (LRT=16.593, p<0.001) were more positive. There was a strong negative shift in reported affective valence as the number of trials increased (LRT=41.588, p=<0.001).

There was no difference in reported valence between the fluctuating reward and fluctuating loss conditions (LRT=1.255, p=0.263). Yet, there was a significant interaction between O_{n-1} (Fig. 3.12: LRT=4.199, p=0.040) and condition; the effect of O_{n-1} was significant in the fluctuating reward condition (LRT=21.216, p<0.001) but not the fluctuating loss condition (LRT=0.644, p=0.422). The interactions between the potential outcome and condition (LRT=0.128, p=0.721) and also between $\bar{R}_{n-1}$ and condition (LRT=0.151, p=0.697) were not significant (Fig. 3.12).

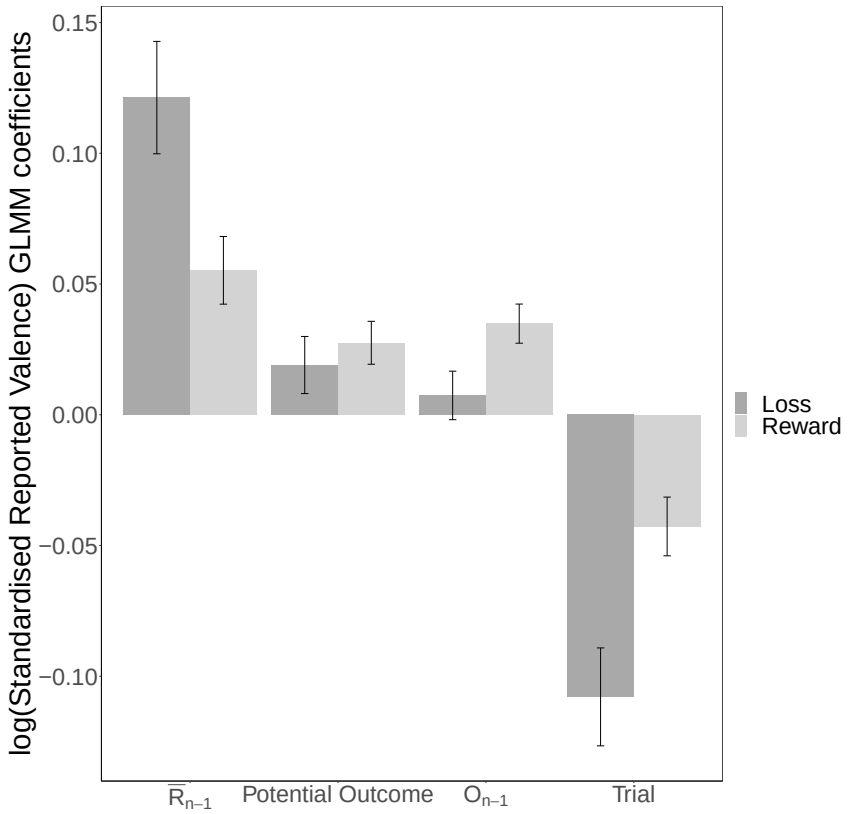


FIGURE 3.12: Standardised GLMM coefficients from the model of the log-transformed affective valence. Error bars represent one standard error

Participants reported significantly greater arousal when $w\text{PE}^2_{n-1}$ (squared weighted prediction error) was higher (LRT=8.516, p=0.004). As with affective valence, the number of trials completed had a strong effect on reported arousal (LRT=12.244, p=<0.001) with participants reporting lower arousal when they had completed a greater number of trials. Neither $\bar{R}_{n-1}$ (average earning rate; LRT=1.570, p=0.210), the potential outcome (LRT=0.105, p=0.746), or condition (LRT=0.013, p=0.909) were significant as main effects. However, there was a significant interaction between the potential outcome and condition (Fig. 3.13: LRT=6.746, p=0.009). Participants tended to report greater arousal when the potential outcome was more positive in the fluctuating reward condition (LRT=3.022, p=0.082) but tended to report greater arousal when the potential outcome was more negative in the fluctuating loss condition (LRT=3.507, p=0.082). The interactions between condition and $\bar{R}_{n-1}$

(LRT=0.005, p=0.942) and condition and $w\text{PE}^2_{n-1}$ (LRT=0.015, p=0.902) were not significant (Fig. 3.13).

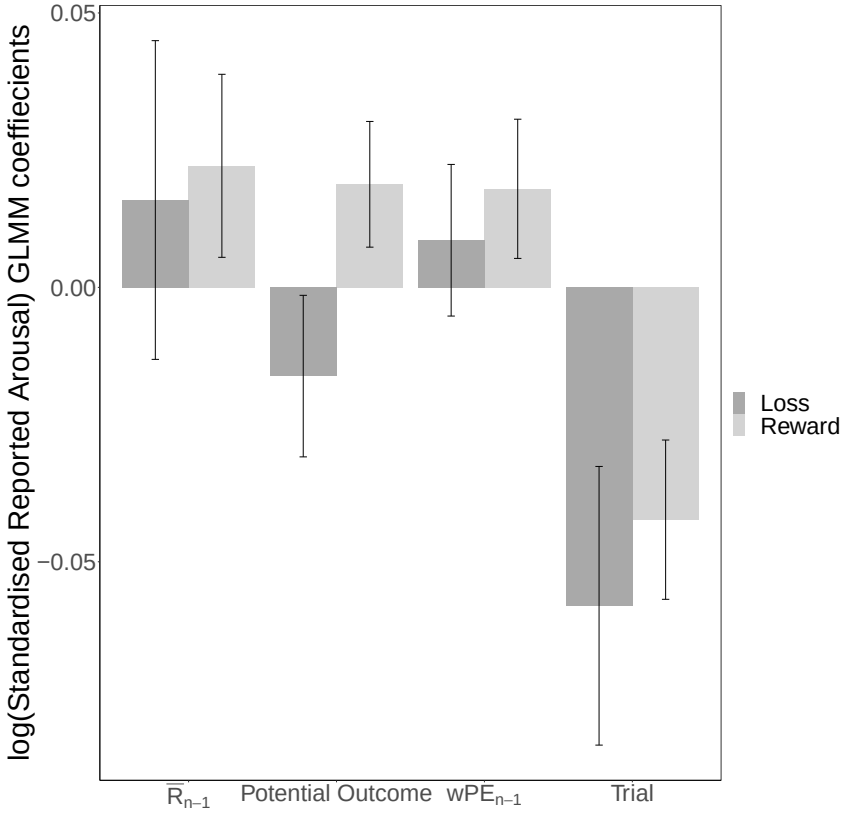


FIGURE 3.13: Standardised GLMM coefficients from the model of the log-transformed reported affective arousal for both the fluctuating reward and fluctuating loss condition. Error bars represent one standard error

3.4 Discussion

The aim of this study was to assess how reward and punisher experience, specifically recent prediction errors, recent squared prediction errors and the average earning rate influence both decision-making within a judgement bias task and self-reported affective state. To achieve this, we used a judgement bias task in which we manipulated reward and punisher experience by systematically varying either the offered punisher or reward and regularly asked participants to report their affective state using an affect grid. Data were analysed using both model-agnostic and a model-dependent analyses, the latter interpreting decision-making on judgement bias task through the lens of a POMDP. This novel model of choice data from a judgement bias task provided a good fit of the data and allowed insight into the influence of reward and punisher experience on decision-making.

3.4.1 Does experience of reward and punisher modulate decision-making?

Reward and punisher experience did influence judgement bias, indicating that past experience is indeed a key determinant of decision-making within the judgement bias task. More specifically, the model-based analysis suggested that the squared weighted prediction error (which is a measure of local uncertainty) was an important factor in the decision to 'stay' or 'go'. Participants were more loss averse, and hence more likely to make the safe ('go') response, when there was greater uncertainty. It has been proposed that individuals should act to reduce uncertainty and minimise surprise (Clark et al., 2018; Friston et al., 2006). In the context of the judgement bias task, this can be achieved by opting for the 'go' response which has a certain outcome (i.e. no reward and no punisher). Hence, participants may have been more risk-averse when the squared weighted prediction error was higher to avoid further uncertainty. However, it should be noted that opting for the risky response rather than the safe response would have provided information about the contingency between the stimulus and outcome and have potentially reduced uncertainty in the longer term.

3.4.2 Does experience of reward and punisher modulate self-reported affective valence?

Participants reported more positive affective valence when the average earning rate was higher. This is in accordance with the theoretical framework for affect proposed by Mendl et al. (2010) in which affective valence is theorised to reflect environmental levels of rewards and punishers; high-reward environments are considered to induce a positively valenced state to drive reward acquisition, and high-punisher environments are suggested to induce a negatively valenced state to promote punisher avoidance. However, it is contrary to more recent suggestions that affect reflects relative but not absolute levels of rewards and punishers; positively valenced affect arises from environments where rewards are greater than expected (Eldar et al., 2016; Rutledge et al., 2014). In contrast to our study, the study conducted by Rutledge et al. (2014) used cumulative earnings as a measure of the absolute favourability of experience within the task, which was not found to influence reported affective valence. However, cumulative earnings may not provide a sufficiently rich measure of current circumstances. For example, an individual could have earned large amounts of money on the first few trials and then lost small amounts of money on every trial thereafter, or could have earned small amounts on all trials, or even lost a large amount initially but won large amounts subsequently, but unlike the average earning rate the cumulative earnings could not distinguish these possibilities. Our results thus support the hypothesis that affective valence should depend on absolute reward and punisher experience. However, as the task is conducted over a very short timescale, further research should be conducted to determine whether absolute levels of rewards and punishers might influence longer-term affect (mood).

Additionally, it is unclear why the average earning rate was not found to influence judgement bias despite influencing affective valence. Likewise, despite influencing decision-making, there was no evidence that the squared prediction error influenced affective valence. This result thus highlights that changes in risk aversion can occur without concurrent changes in self-reported affective valence, and vice versa. This should be explored further in future studies.

However, participants did report greater arousal when recent outcomes had been less predictable. This perhaps reflects the fact that surprise is generally considered to be a high arousal state (Fontaine et al., 2007; Reisenzein, 1994), which could potentially function to prepare individuals to act to reduce future uncertainty.

3.4.3 How do variables underlying decision-making behaviour relate to affective valence and arousal?

The extent to which the squared weighted prediction error (indicating less predictable/more surprising outcomes in recent trials) modulated a participant's decision to stay or go in the judgement bias task depended on affective valence. Specifically, participants whose decision-making was modulated to a lesser extent by surprising outcomes reported overall more negative affective valence. Depression has been associated with blunted prediction error signalling (Gradin et al., 2011; Kumar et al., 2008; Steele et al., 2004). Thus, although this study was conducted in a non-clinical population, this finding may further support the relationship between affective state and prediction error responsivity.

This result is also reminiscent of cautious optimism; the finding that positive affect can induce greater caution, despite a more optimistic belief about the outcome of decisions (Isen et al., 1988; Isen and Patrick, 1983; Nygren et al., 1996). Cautious optimism has been explained as a self-protecting mechanism which leads individuals to make decisions that allow them to maintain their positive affect, specifically by altering the value of the loss (Isen et al., 1988; Isen and Patrick, 1983; Nygren et al., 1996). Therefore, an alternative or additional explanation for the finding that affective valence correlates with the model parameter characterising the influence of uncertainty on loss aversion is that individuals in more positive affective states were more averse to losing money and spoiling their positive affective state in periods of greater surprise or uncertainty than those in less positive affective states.

Given that we cannot infer causation in this instance, it is also possible that participants that had been less cautious following less predictable outcomes were in a more negatively valenced affective state resulting directly from their lack of caution (i.e. greater losses due to greater risk-seeking behaviour). However, although weaker loss aversion could lead to greater losses it could also lead to greater gains.

Loss sensitivity was found to be associated with reported arousal; individuals who were less sensitive to losses reported overall greater arousal. This might suggest that individuals who were more willing to take risks (via a decreased value of loss) were also in a state of greater arousal, perhaps induced by risk-taking. This would be supported

by the finding that risking-taking increases physiological measure of arousal (Krueger et al., 2005; Meyer et al., 2000; Wulfert et al., 2005).

3.4.4 Does reward and punisher experience influence trial initiation latency?

Although the main aim of the study was to investigate decision-making, the trial initiation latency provided an opportunity to investigate another aspect of behaviour likely to be influenced by prior experience, namely vigour. Vigour has operationally been defined as inverse latency; it corresponds to the alacrity with which actions are made. Vigour can be optimised in accordance with the opportunity cost of time; actions should be made more vigorously when the rewards forgone by acting slothfully are greater (Niv et al., 2007). The opportunity cost of time is determined by the overall reward rate, encouraging participants to capitalise on periods of good fortune and delay unfavourable outcomes. Participants were indeed faster to initiate trials when the average earning rate was higher which is consistent with previous findings (Griffiths and Beierholm, 2017; Guitart-Masip et al., 2011; Niv et al., 2007).

The weighted prediction error was also a determinant of trial initiation but in the opposite direction. Participants were slower to initiate trials following unexpectedly good outcomes compared with unexpectedly bad outcomes. One potential explanation for this is that negative prediction errors might imbue the individual with a sense of urgency to improve their situation, whereas positive prediction errors might inform the individual that no behavioural change is necessary, or that expectations could be met with reduced effort.

The average earning rate and weighted prediction error therefore have opposing effects on vigour; participants are faster when both the average earning rate is higher and weighted prediction error is more negative. This might reflect differences in timescale; the average earning rate being generated over a longer timescale than the weighted prediction error. Experiencing unexpectedly aversive outcomes might invigorate the individual, allowing them to move on to the next trial where they might recoup their loss more quickly, whereas longer-term poor conditions might promote sloth (i.e. encourage slower initiation) to delay trials where a poor outcome is expected. Similarly, an individual might be slower to savour a positive outcome especially if the following outcome may not be as good, but when the environment is particularly rewarding it is advantageous to capitalise on those good conditions and initiate trials with greater vigour.

While significant across both conditions, the effect of both the prediction error and average earning rate was more pronounced in the fluctuating reward condition. This might reflect that loss avoidance can be guaranteed on any trial by making the safe response, but reward gain cannot similarly be assured. So, the average earning rate prediction error and more relevant in the context of variable rewards. This might also provide an explanation as to why the influence of surprise on judgement bias was greater in the reward condition.

3.4.5 How do immediate rewarding and punishing experiences influence decision-making and affect?

Although we were particularly interested in the effects of prediction error, squared prediction error, and average earning rate, as we considered that these aspects of reward and punisher experience would be most relevant to future experience and hence behaviour, it is important to also consider also the impact of the magnitude of the offered and most recent rewards and losses on self-reported affect and decision-making.

Individuals were more likely to make the risky response when the offered reward was higher or potential loss was lower, indicating that they were attending to information about the outcomes and acted rationally (Von Neumann and Morgenstern, 1953). Additionally, participants reported more positive affective valence when the most recent trial had offered a higher reward or lower loss, which could reflect that the opportunity to win greater amounts is likely to induce a positive emotional state, while the potential to lose greater amounts is likely to induce a negative emotional state (Rolls, 2013). Affective arousal was also influenced by the potential outcomes. Greater affective arousal tended to be reported when both the offered reward and threatened loss were higher, demonstrating that high stakes trials required greater alertness.

Participants were slower to initiate trials when the potential reward was higher in the fluctuating reward condition. We would expect this effect to be in the opposite direction, with higher rewards leading to greater invigoration. Our sole explanation for this is that this reflects some form of speed/accuracy trade-off whereby the individual takes time to prepare for the trial. Participants also tended to be slower to initiate trials when the potential loss was higher in the fluctuating loss condition, perhaps reflecting an attempt to delay more unfavourable trials or potentially also an increased preparation time.

A more favourable outcome on the previous trial led to more positive reported affective valence, although only in the fluctuating reward condition. We found that the outcome on the most recent trial provided a better explanation of variation in affective valence and arousal than recent prediction errors or recent squared prediction errors. A study by Rutledge et al. (2014) found that more positive prediction errors were associated with increased happiness. However, they reported high forgetting factors in the mapping of recent prediction errors to affective valence, and so there might have been, at least in some participants, a similarly outsized effect of just the last trial.

The task itself influenced reported affect and decision-making, potentially due to its repetitive nature. Participants reported more negative affective valence and arousal as the task progressed which is indicative of boredom. They also became less risky in their decision-making which could reflect boredom (negative affect) induced risk-aversion. Despite reporting lower arousal, participants were faster to initiate trials as the test session progressed. This could reflect either being more familiar with the

task or also an attempt to get through the task more quickly in the face of increasing boredom.

3.5 Conclusions (see Fig. 3.14)

This study revealed a number of novel relationships between reward experience, affective state, and behaviour. It implicates the average earning rate in affective valence, with individuals reporting more positive affective valence when their average earning rate was higher. The study also highlights the role of predictability in decision-making, with increased risk aversion when recent outcomes were more unpredictable/surprising. Moreover, the extent to which predictability modulated decision-making was associated with self-reported affective valence; more negatively valenced affect was linked to weaker predictability-dependent modulation of decision-making. Additionally, we found that unexpectedly good outcomes decreased vigour to initiate trials, while unexpectedly poor outcomes increased vigour to initiate trials. These findings would not have been revealed without the development of a novel computational model for examining decision-making on the judgement bias task, which also demonstrated that variation in loss sensitivity determined variation in judgement bias. The study also confirmed many well-known relationships, such as that between a measure of vigour and the average earning rate and between potential outcomes and risk-taking. Overall, the results provide evidence that rewarding and punishing experiences determine both future decision-making and subjectively experienced affective state.

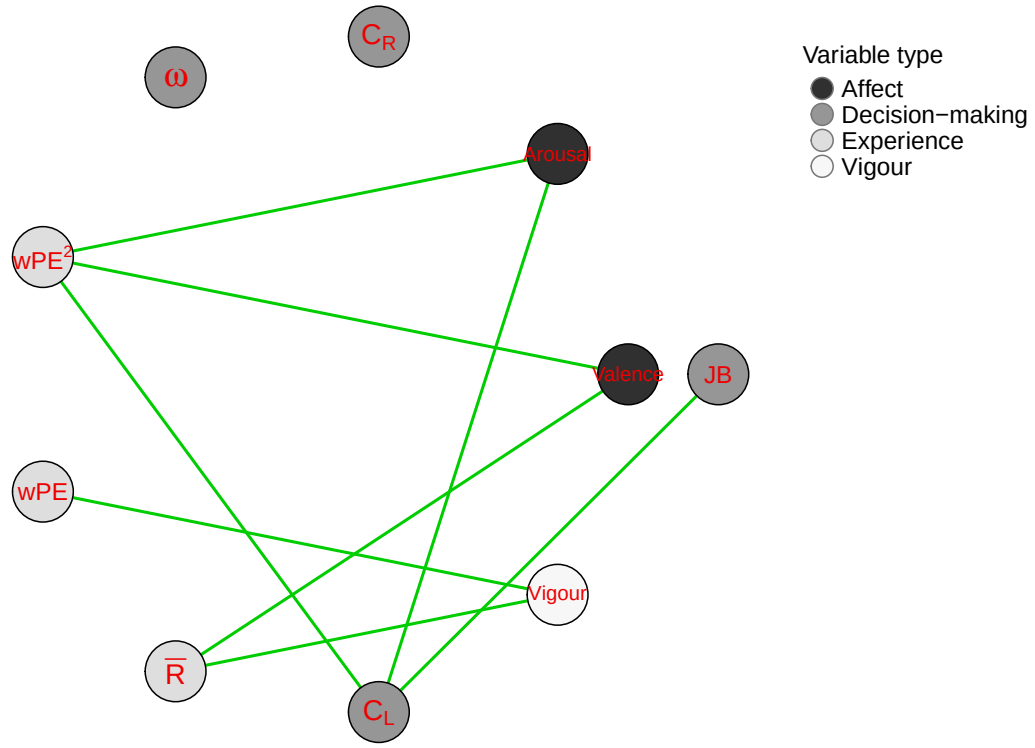


FIGURE 3.14: Diagrammatic summary of the results of Chapter 3: the nodes represent variables relating to affect, decision-making (where C_L denotes loss/punisher sensitivity, C_R denotes reward sensitivity, ω denotes the prior belief about outcomes, and JB denotes judgement bias), experience (where wPE denotes the weighted reward prediction error, $w\text{PE}^2$ denotes the squared weighted reward prediction error, and $\bar{R}$ denotes the average earning/reward rate), and vigour, the lines between nodes represent observed associations between those variables.

Chapter 4

How do fluctuating primary rewards influence human affect and judgement bias?

Chapter summary: The translatability of behavioural measures of affect between human and non-human animals is important to the accurate assessment of animal welfare and for the development of novel pharmacological treatments for affective disorders. However, there is a disparity between the modality of reinforcers used in studies with human and non-human animals; human studies typically use secondary reinforcers (e.g. money) while non-human animal studies typically use primary reinforcers (e.g. food and electric footshocks). To address this deficiency, we repeated our human judgement bias task (Chapter 3) using apple juice as a reward, and cold salty tea as a punisher, instead of monetary rewards and losses. The potential volume of apple juice on each trial fluctuated according to a noisy sine wave, and computational modelling of the judgement bias data allowed insight into the cognitive processes underlying decision-making. The results were largely similar to that reported in Chapter 3: the average earning rate determined affect and vigour, unpredictability influenced judgement bias (although via reward sensitivity and altered prior beliefs as opposed to loss sensitivity in Chapter 3), and affective valence was correlated with the parameters characterising the extent to which unpredictability modulated judgement bias. Thus, we have developed a task that more closely resembles the judgement bias tasks used with non-human animals, and with results that replicate and hence support the reliability and validity of the results of the human monetary task.

4.1 Introduction

Typically, operant testing in non-human animals is conducted using primary reinforcers such as food or electric shocks, whereas operant testing in humans involves secondary reinforcers such as monetary gain or loss. However, a number of studies have identified differences in the neural processing of primary and second reinforcers (Beck et al., 2010; Delgado et al., 2011; Grimm and See, 2000; Sescousse et al., 2013). Consequently, making comparisons and drawing conclusions between the results of human and non-human animal tasks is complicated by the issue that we cannot preclude the possibility that any differences may relate to reinforcer type (primary vs. secondary) as opposed to species differences.

The need for translatable results is particularly apparent in the study of affect in non-human animals. Measures of animal welfare are commonly derived from observations in humans, for example anhedonia in human depression (American Psychiatric Association, 2013) or cognitive biases first described in humans (MacLeod et al., 1986; Williams et al., 1996; Wright and Bower, 1992), and the development of new treatments for affective disorders in humans depends on the use of animal models of affect (Rupniak, 2003). Judgement bias has been demonstrated to overall provide a measure of affective states that are pharmacologically induced in a range of non-human animals (Chapter 2). However, assessing the extent to which judgement bias corresponds to subjective reports of affect in humans would enhance its value as measure of affect and further support its validity. Moreover, studies with humans may help to provide a better insight into precisely what judgement bias measures in terms of cognition, and how reward and punisher experience alters decision-making. While studies have examined whether measures of subjectively reported affect in humans correlate with judgement bias, and have indeed found that risk aversion correlates with PANAS (Positive and Negative Affective Scale) questionnaire results (Paul et al., 2011) and state anxiety (Anderson et al., 2012), these studies used secondary reinforcers and hence may not be entirely comparable to the studies in non-human animals which use primary reinforcers.

To address this deficiency, we aimed to develop a translation of the automated rat judgement bias task (Jones et al., 2018) for humans which uses primary reinforcers: apple juice and cold salty tea. Using this novel variant of the judgement bias task alongside computational modelling, we aimed to examine how reward and punisher experience (specifically recent prediction errors and the average earning rate) relates to judgement bias, affect, and vigour, and how this compares to a monetary version of the task (see Chapter 3) in which the average earning rate was a key determinant of affective valence, while surprise was a key determinant of decision-making.

4.2 Methods

Participants

Thirty-three people (26 female, 7 male, mean age±SE=31.182±2.0139) from the Bristol Veterinary School community participated in the study. The inclusion criteria were that the participant enjoyed apple juice, was not allergic to apple juice, salt, or black tea, was not hypertensive or had any medical condition that meant that they must limit their salt intake, and was over the age of 18. Participants were asked to abstain from drinking anything (except water) or eating in the hour prior to the study. To cover their time and expenses, participants were paid £7 for the hour-long session. To encourage full engagement with the study, participants were informed that the top-three ranking participants in terms of accuracy would receive an additional £7 bonus. All participants provided written informed consent, and the study was approved by the Faculty of Science Research Ethics Committee at the University of Bristol.

Apparatus (see Fig. 4.1)

The task was conducted on a laptop (Dell Latitude) which was connected to two syringe pumps (World Precision Instruments, Hertfordshire, UK; SPLG100) which were set to pump liquid at a rate of 2ml per minute. Sterilised food grade PVC tubing (OD:11 mm, ID: 8mm) attached to syringes (Becton Dickinson, Swindon, UK; 50ml Plastipak) driven by the syringe pumps were connected via a tube connector to smaller sterilised PVC tubing (OD: 6mm, ID: 4mm) and held in place in front of the participant using a retort clamp and stand, the height of which could adjusted by the participant at the start of the experiment. Individuals made responses on a keyboard connected to the laptop. The task was written in MATLAB (MathWorks, Natwick, MA, USA) using the PsychToolBox extensions.

FIGURE 4.1: Photograph of the experimental set-up. The two syringe pumps containing salty tea or apple juice are located on the left side of the table, the laptop is located on the back centre of the table, a retort stand and clamp holding the tubes delivering the juice are located in front of the laptop, and a keypad is also located in front of the laptop.

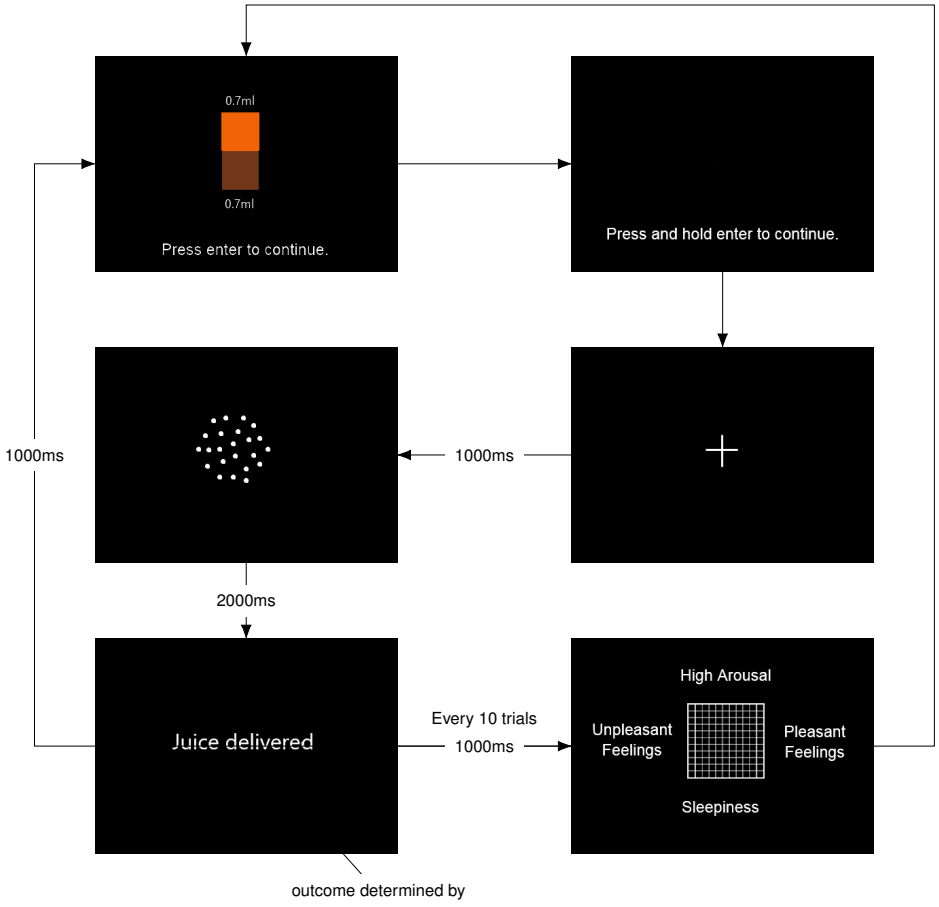
4.2.1 Training

The methodology was adapted from that of the human monetary judgement bias task (Chapter 3). On each trial of the task, participants were instructed to press and hold the enter key before being presented with a fixation cross for 1000ms followed by a random dot kinematogram (RDK) displayed for 2000ms, which across trials varied in direction of motion (leftwards or rightwards) and coherence (proportion of dots moving in a coherent direction: 0.01, 0.02, 0.16, or 0.32). Participants had two options when the RDK was presented: to release the key prior to the end of the two second RDK presentation ('go') or continue to hold the key for two seconds ('stay'). The outcome associated with either response depended on the stage of training and the direction of the RDK with one direction being favourable and the other being unfavourable (either leftwards or rightwards, counterbalanced across participants).

Training comprised two practice blocks of 48 trials. The aim of the first practice block was to introduce the participants to the task and train them on the correct responses to the RDK. In this first practice block, the word 'correct' was shown on screen for correct responses; 'staying' when the RDK moves in the favourable direction and 'going' otherwise. The direction of motion of the RDK was very easy to detect (coherence=0.32) on all trials, and the direction of motion was leftwards on 50% of trials and rightwards on the remaining 50%. The aim of the second practice block was to acquaint participants with the delivery and taste of the apple juice (Morrisons, West Yorkshire, UK; Apple Juice from Concentrate) and salty tea. Salty tea was prepared each morning following Pauli et al. (2015) with two black tea bags (Bettys & Taylors of Harrogate, Harrogate, UK; Yorkshire Tea) and 29g of salt per litre of boiling water which was chilled prior to data collection. In this block, the direction of the RDK was easy to detect (coherence=0.16) on 50% of trials, of which half were leftwards and half rightwards, and moderately difficult to detect (coherence=0.02) on the remaining 50%, of which half were leftwards and half rightwards. The volume of juice received for stay responses when the direction of the RDK was favourable was 0.7ml and likewise the volume of salty tea received for stay responses when the RDK was unfavourable was 0.7ml. While juice was delivered and for 3000ms following delivery, the words, 'Juice delivered' were displayed on screen while during salty tea delivery and for 3000ms following this, the words 'Salty tea delivered' were displayed on screen. The potential volume of juice on each trial (i.e. '0.7ml') was shown above an orange coloured bar with a height proportional to the potential juice volume, similarly the volume of salty tea (i.e. '0.7ml') was displayed above a brown bar positioned directly below the orange coloured bar with a height proportional to the potential tea volume. These bars were shown prior to the instruction to press and hold the enter key. The participant received nothing for making the go response and the words 'Nothing delivered' were displayed on screen for 3000ms. Across all blocks the direction and coherence level of the random dot kinematogram was randomised prior to the start of the study and was identical for all participants.

4.2.2 Testing (see Fig. 4.2)

The test session comprised 180 trials and on 50% of trials the direction of the RDK was leftwards and on the remaining 50% it was rightwards, while the coherence level was either low (0.01), moderate (0.02), or high (0.16). Hence each possible stimulus was shown on 16.7% of trials. The potential volume of juice fluctuated across trials according to a noisy sine wave with a mean volume of 0.686ml and standard deviation of 0.325ml, ranging from a minimum of 0.031ml to a maximum of 1.171ml (Fig. 4.3). The potential volume of salty tea remained at 0.7ml throughout testing. As in the second practice block, the potential volume of juice and salty tea was displayed on screen both as text and graphically using two coloured bars with heights proportional to the volume of juice and salty tea. Prior to the first trial and every subsequent 10 trials, participants were asked to report how they were feeling using a computerised affect grid (Killgore, 1998). To complete the affect grid, participants had to move a cross which was initially central in the grid to the location that best described their current mood using the arrow keys on a keyboard. Horizontal movements represented changes in mood valence, with movements to the left reporting a more positively valenced mood. Vertical movements represented arousal, with upwards movement reporting higher levels of arousal.



	Stay (hold key)	Go (release key)
Favourable direction	Juice	Nothing
Unfavourable direction	0.7ml salty tea	Nothing

FIGURE 4.2: Structure of the human primary reward judgement bias test session

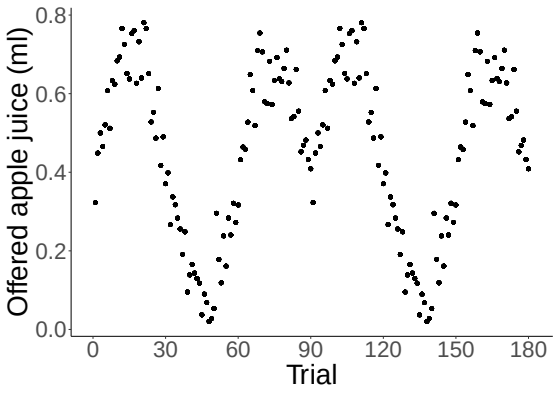


FIGURE 4.3: Offered volume of apple juice on each trial of the judgement bias task

4.2.3 Model dependent analysis

Judgement bias choice data ('stay' – risky, or 'go' – safe) were fitted to the partially observable Markov decision process model used to model human judgement bias in Chapter 3. Briefly, we consider that participants transition through a two-dimensional state space in which they construct and maintain a belief about the direction of the RDK that is informed by their momentary observations of the RDK across time. The probability that an individual will opt for the safe response will depend on their likely transitions through the state space and their subjective value of occupying each state. In addition to the direction and coherence of the presented RDK, this is determined by a number of parameters including σ, which reflects the ability of the participant to discriminate between the stimuli, a lapse rate (λ), and a set of parameters which characterise the influence of reward and punisher experience, including the average earning rate ($\bar{R}_{n-1}$), weighted prediction error ($w\text{PE}_{n-1}$), squared weighted prediction error ($w\text{PE}^2_{n-1}$), and most recent outcome (O_{n-1}) on reward sensitivity (C_R) which scales the true volume of juice, punisher sensitivity (C_P) which scales the true volume of salty tea, and the prior belief that the trial will be rewarded (ω) which allows the delivery of juice to be perceived as more or less probable. The model assumes that $\bar{R}_{n-1}$, $w\text{PE}_{n-1}$, $w\text{PE}^2_{n-1}$, and O_{n-1} influence C_R, C_P, and ω additively, with beta values which determine the extent of their influence as follows:

$$\omega = e^{\left(\beta_0^{\omega} + \beta_{\bar{R}}^{\omega}\bar{R}_{n-1} + \beta_{w\text{PE}}^{\omega} w\text{PE}_{n-1} + \beta_{w\text{PE}^2}^{\omega} w\text{PE}^2_{n-1} + \beta_O^{\omega} O_{n-1}\right)} \tag{4.1}$$

$$C_R = e^{\left(\beta_0^{C_R} + \beta_{\bar{R}}^{C_R}\bar{R}_{n-1} + \beta_{w\text{PE}}^{C_R} w\text{PE}_{n-1} + \beta_{w\text{PE}^2}^{C_R} w\text{PE}^2_{n-1} + \beta_O^{C_R} O_{n-1}\right)} \tag{4.2}$$

$$C_P = e^{\left(\beta_0^{C_P} + \beta_{\bar{R}}^{C_P}\bar{R}_{n-1} + \beta_{w\text{PE}}^{C_P} w\text{PE}_{n-1} + \beta_{w\text{PE}_n^2}^{C_P} w\text{PE}^2_{n-1} + \beta_O^{C_P} O_{n-1}\right)} \tag{4.3}$$

A learning rate ($\alpha_{\bar{R}}$) which determines the speed of learning following reward and punisher experience and hence modulates the average earning rate and forgetting factor ($\gamma_{w\text{PE}}$) which allows variation in how quickly recent prediction errors are forgotten are also included.

We assessed whether the prediction error between the prediction either prior to or following RDK onset and the outcome was most relevant to decision-making. Consistent with the monetary task, the prediction error most relevant to decision-making was that between the prediction prior to stimulus onset and the outcome, as opposed to between the prediction post-stimulus and outcome (ΔAIC=516.063, comparing all models in which decision-making depended on wPE using either the pre or post stimulus onset prediction error).

The aim of the model-dependent analysis was to investigate the relationship between primary reward and punisher experience and decision-making, specifically to examine which, if any, cognitive processes involved in decision-making were modulated by different aspects of reward and punisher experience. Hence, analysis of the parameter estimates from the POMDP model examined the key predictions of this study.

Potential models were fitted in a stepwise manner, first assessing whether σ and λ both improved model fit, then assessing the contribution of each aspect of reward and punisher experience individually, before finally assessing the goodness of model fit when parameters determining reward sensitivity, punisher sensitivity, and the prior belief were jointly included. Model comparison was made using the AIC (Akaike's information criterion) values. The parameter estimates of interest from the most parsimonious (lowest AIC) model were analysed using permutation tests (PT) to assess whether they varied significantly from zero, and using general linear models which included and mean reported affective valence and mean reported affective arousal (from each individual's set of affect grids) as dependent variables to assess how the parameter estimates corresponded to reported affect. Model-fitting was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bris.ac.uk/acrc/, as described in Chapter 3.

4.2.4 Model-agnostic statistical analysis

Judgement bias choice, log-transformed trial initiation latency, log-transformed reported valence (affect grid x-coordinate), log-transformed reported arousal (affect grid y-coordinate) were analysed using general linear mixed models (GLMM) with a random effect of participant in R (R Core Team, 2017) using the lme4 (Bates et al., 2015) and nlme (Pinheiro et al., 2018) packages. The judgement bias choice model included a binomial link function while the remaining models included a Gaussian link function. The model assumptions were verified in all cases, where data is log-transformed this transformation allowed the model assumptions to be met. All models included the potential outcome ($\frac{\text{offered juice (ml)}-\text{potential salty tea (ml)}}{2}$) and the number of trials

completed prior to the current trial as independent variables. The GLMM of judgement bias also included a variable for the stimulus presented (combining coherence and direction), and the models of trial initiation latency and reported affect included variables encompassing reward and punisher experience which were extracted from the best POMDP model (i.e. $\bar{R}_{n-1}$, $w\mathrm{PE}_{n-1}$, $w\mathrm{PE}^2_{n-1}$, and O_{n-1}).

Analysis of the affect grid data allowed us to examine the relationship between primary reward and punisher experience and affect; a key aim of this chapter. Analysis of the trial initiation data allowed us to examine whether primary reward and punisher experience altered vigour. These model-agnostic analyses also allowed us to assess whether there were time-dependent changes in affect, vigour, and decision-making; and to assess the extent to which participants were attending to the potential outcome of each trial.

Due to correlations between $w\mathrm{PE}_{n-1}$, $w\mathrm{PE}^2_{n-1}$, and O_{n-1}, the predictor variables were not included in the same GLMM but instead were included separately and each model was then compared using their AIC values. The GLMM which provided the best fit was analysed further. In the model of trial initiation latency this was $w\mathrm{PE}^2_{n-1}$ (ΔAIC=4.775; comparing saturated model containing $w\mathrm{PE}^2_{n-1}$ and not $w\mathrm{PE}_{n-1}$ or O_{n-1} with the next best-fitting saturated model which contained O_{n-1} but not $w\mathrm{PE}_{n-1}$ or $w\mathrm{PE}^2_{n-1}$), in the model of reported valence this was O_{n-1} (ΔAIC=1.214; comparing saturated model containing O_{n-1} and not $w\mathrm{PE}_{n-1}$ or $w\mathrm{PE}^2_{n-1}$ with the next best-fitting saturated model which contained $w\mathrm{PE}_{n-1}$ but not O_{n-1} and not $w\mathrm{PE}^2_{n-1}$), and in the model of reported arousal this was $w\mathrm{PE}^2_{n-1}$ (ΔAIC=0.241; comparing saturated model containing $w\mathrm{PE}^2_{n-1}$ and not $w\mathrm{PE}_{n-1}$ or O_{n-1} with the next best-fitting saturated model which contained $w\mathrm{PE}_{n-1}$ but not O_{n-1} or $w\mathrm{PE}^2_{n-1}$). Likelihood ratio tests were used to assess the significance of each dependent variable.

4.3 Results

4.3.1 Model dependent results

The model of judgement bias with the lowest AIC value included the following parameters: σ, λ, β^{ω}_{0}, $\beta^{\omega}_{w\mathrm{PE}^2_n}$, $\beta^{C_R}_{0}$, $\beta^{C_R}_{w\mathrm{PE}^2_n}$ (Table 4.1, Fig. 4.4). The model-derived probability of making a 'stay' response was found to be a strongly significant predictor of the observed response when analysed using a binomial GLMM with a random effect of subject (Fig. 4.5: LRT=2087, $p<0.001$). Inclusion of a fixed learning rate ($\alpha_{\bar{R}}$; ΔAIC=49.012, comparing all models in which decision-making depended on $\bar{R}$ and $\alpha_{\bar{R}}$ was allowed to vary, to the same set of models in which $\alpha_{\bar{R}}$ was fixed) and fixed forgetting factor provided a more parsimonious fit ($\gamma_{w\mathrm{PE}}$; ΔAIC=195.098, comparing all models in which decision-making depended on wPE and $\gamma_{w\mathrm{PE}}$ was allowed to vary, to the same set of models in which $\gamma_{w\mathrm{PE}}$ was fixed) than models in which the learning rate and forgetting factor fitted to each individual's data was allowed to vary.

TABLE 4.1: ΔAIC values for computational models of judgement bias choice data: the difference between the AIC values of each model and the AIC value for the best model

Model parameters	**ΔAIC**
σ, λ, β_0^{ω}, $\beta_{w\mathrm{PE}_n^2}^{\omega}$, $\beta_0^{C_R}$, $\beta_{w\mathrm{PE}_n^2}^{C_R}$	0.000
σ, λ, $\beta_0^{C_R}$, $\beta_{w\mathrm{PE}_n^2}^{C_R}$, $\beta_0^{C_P}$, $\beta_{w\mathrm{PE}_n^2}^{C_P}$	36.914
σ, λ, $\beta_0^{C_R}$, $\beta_{w\mathrm{PE}_n^2}^{C_R}$	131.254
σ, λ, $\beta_0^{C_R}$, $\beta_{\bar{R}}^{C_R}$	142.912
σ, λ, $\beta_0^{C_R}$	187.258
σ, λ, β_0^{ω}, $\beta_{w\mathrm{PE}_n^2}^{\omega}$, $\beta_0^{C_P}$, $\beta_{w\mathrm{PE}_n^2}^{C_P}$	193.228
σ, λ, $\beta_0^{C_P}$, $\beta_{w\mathrm{PE}_n^2}^{C_P}$	197.383
σ, λ, $\beta_0^{C_R}$, $\beta_{O_{n-1}}^{C_R}$	204.275
σ, λ, $\beta_0^{C_P}$	211.623
σ, λ, $\beta_0^{C_P}$, $\beta_{\bar{R}}^{C_P}$	214.939
σ, λ, $\beta_0^{C_P}$,$\beta_{w\mathrm{PE}_n}^{C_P}$	224.203
σ, λ, $\beta_0^{C_R}$, $\beta_{\bar{R}}^{C_R}$, $\beta_{w\mathrm{PE}_n^2}^{C_R}$	231.082
σ, λ, β_0^{ω}, $\beta_{w\mathrm{PE}_n^2}^{\omega}$	254.435
σ, λ, β_0^{ω},$\beta_{w\mathrm{PE}_n}^{\omega}$	255.159
σ, λ, $\beta_0^{C_P}$, $\beta_{O_{n-1}}^{C_P}$	262.486
σ, λ, $\beta_0^{C_R}$, $\beta_{w\mathrm{PE}_n}^{C_R}$	282.635
σ, λ, β_0^{ω}, $\beta_{O_{n-1}}^{\omega}$	344.424
σ, λ, β_0^{ω}	438.609
σ, λ, β_0^{ω}, $\beta_{\bar{R}}^{\omega}$	463.960
σ, λ	1023.645
σ	1214.208
λ	1412.778

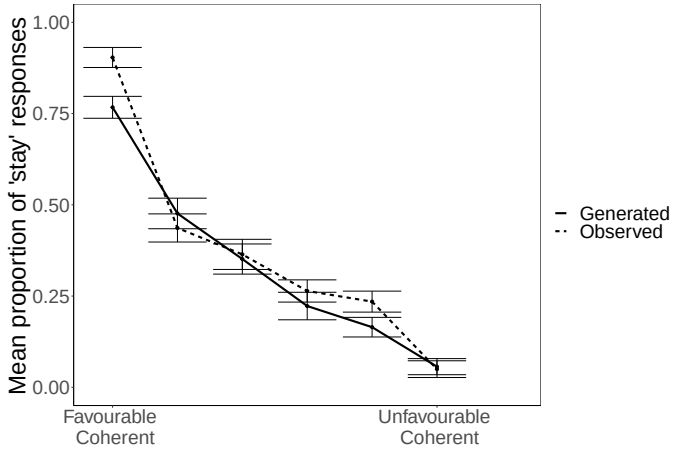


FIGURE 4.4: The mean proportion of 'stay' responses from the model-generated and observed judgement bias data. Errors bars represent one standard error

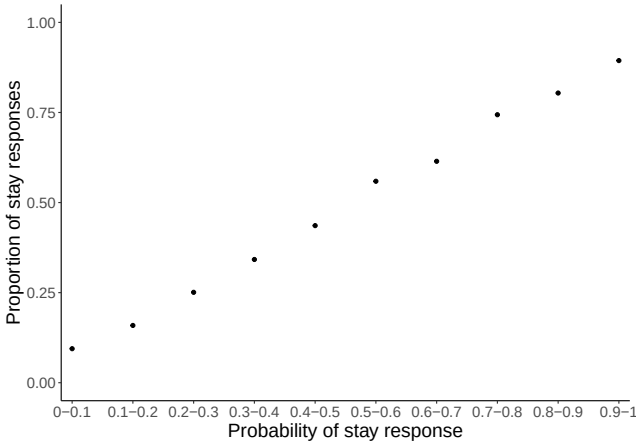


FIGURE 4.5: The proportion of 'stay' responses for intervals of model-derived probabilities of executing the 'stay' response

The estimates of β_0^{ω} (the baseline prior probability) were significantly lower than zero (PT: mean±SE=-1.378±0.493, p=0.007) while the estimates of $\beta_0^{C_R}$ (the baseline reward sensitivity) were significantly higher than zero (PT: mean±SE=1.964±0.293, p<0.001) indicating that participants both had an increased value of the juice relative to its absolute value (i.e. volume) but had a stronger belief overall that each trial would be punished. Participants were less sensitive to rewards when $w\mathrm{PE}^2$ was higher (PT, $\beta^{C_R}_{w\mathrm{PE}^2_n}$ – characterising the effect of the squared prediction error on reward sensitivity: mean±SE=-0.316±0.109, p=0.003) although the estimates of $\beta^{\omega}_{w\mathrm{PE}^2_n}$ (characterising the effect of the squared prediction error on the prior probability) did not differ significantly from zero (p=0.827).

More negative reported valence was associated with increased reward sensitivity ($\beta_0^{C_R}$: LRT=5.556, p=0.018), weaker modulation of reward sensitivity by $w\mathrm{PE}^2_{n-1}$ ($\beta^{C_R}_{w\mathrm{PE}^2_n}$: LRT=10.511, p=0.001), and more negative values of $\beta^{\omega}_{w\mathrm{PE}^2_n}$ (LRT=6.789, p=0.009) where more negative values indicate a decreased belief that a trial would be rewarded as $w\mathrm{PE}^2_{n-1}$ increased. There was no significant correlation between β_0^{ω} and reported valence (LRT=0.164, p=0.686). The estimates of β_0^{ω} (LRT<0.001, p=0.997), $\beta^{\omega}_{w\mathrm{PE}^2_n}$ (LRT=0.311, p=0.577), $\beta_0^{C_R}$ (LRT=1.941, p=0.164), and $\beta^{C_R}_{w\mathrm{PE}^2_n}$ (LRT=0.195, p=0.659) did not depend significantly on reported arousal.

4.3.2 Model-agnostic statistical analyses

Judgement bias

The decision to 'stay' (risky) or 'go' (safe) depended on the stimulus presented (LRT=2154.599, p<0.001), with more 'go' responses made towards RDKs that were unfavourable and increasingly coherent, and the potential outcome (Fig. 4.6: LRT=8.543, p=0.003), with a greater likelihood of the 'go' response when the potential juice volume was lower. Participants became increasingly risk-averse as they completed more trials (LRT=27.733, p<0.001).

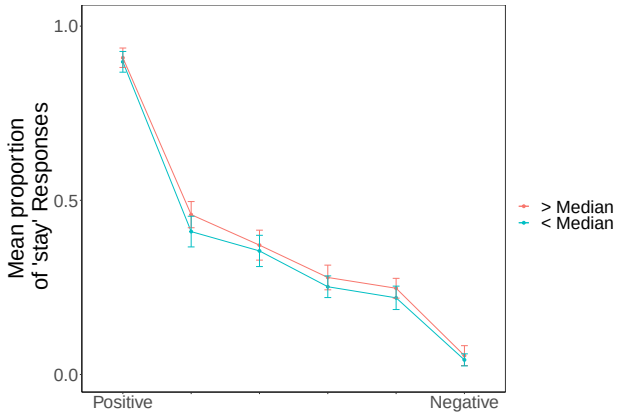


FIGURE 4.6: The mean proportion of 'stay' responses for each stimulus split by whether the expected value was higher (red) or lower (blue) than the median value

Affective state

More positive values of $\bar{R}_{n-1}$ (the average earning rate, Fig. 4.7: LRT=8.019, p=0.005), a greater volume of juice offered on the previous trial (Fig. 4.7: LRT=15.353, p<0.001), and more positive values of O_{n-1} (the previous outcome, Fig. 4.7: LRT=4.302, p=0.038) were significantly associated with more positive affective valence as reported using the affect grid. However, reported affective valence became more negative as the number of trials completed increased (Fig. 4.7: LRT=6.608, p=0.010).

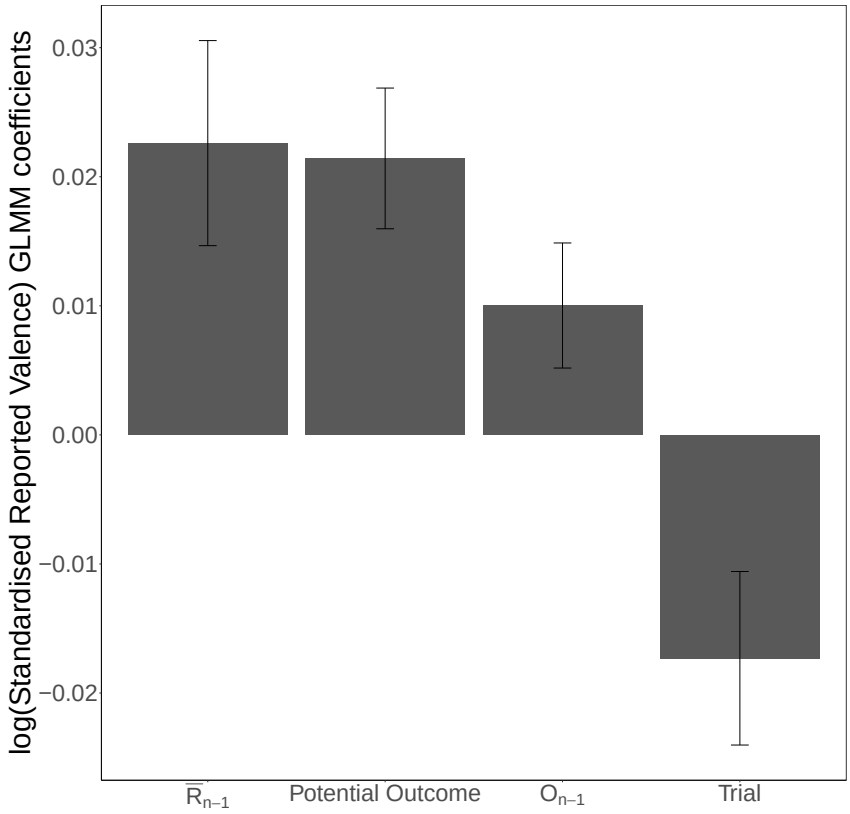


FIGURE 4.7: Standardised GLMM coefficients from the model of the log-transformed reported valence. Error bars represent one standard error

Participants tended to report greater arousal using the affect grid when the potential outcome (Fig. 4.8: LRT=2.767, p=0.096) on the most recent trial was higher, while $w\text{PE}^2_{n-1}$ (squared weighted prediction error, Fig. 4.8: LRT=0.261, p=0.610) and $\bar{R}_{n-1}$ (Fig. 4.8: LRT=0.131, p=0.717) were not found to significantly influence reported arousal. Reported arousal decreased significantly throughout the test session (Fig. 4.8: LRT=11.702, p=0.001).

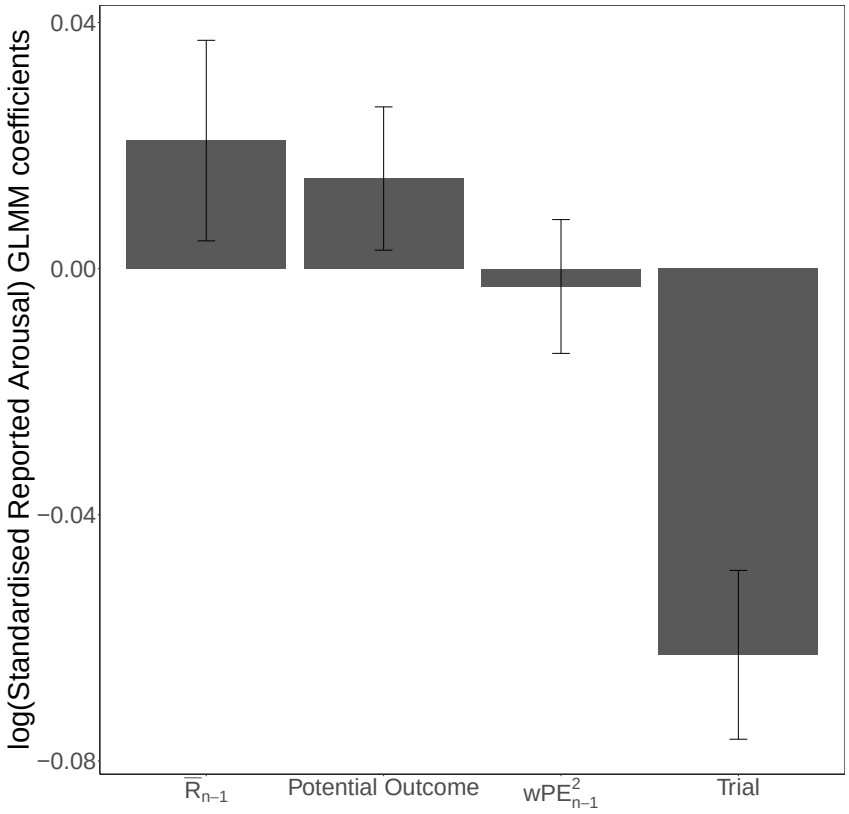


FIGURE 4.8: Standardised GLMM coefficients from the model of the log-transformed reported arousal. Error bars represent one standard error

Trial initiation

All aspects of reward and punisher experience influenced the participants' latency to initiate trials. Participants were slower to initiate trials both when the volume of juice offered on the trial was higher (Fig. 4.9: LRT=5.091, p=0.024) and when $w\text{PE}^2_{n-1}$ (squared weighted prediction error) was greater (Fig. 4.9: LRT=18.395, p<0.001), while they were faster to initiate trials when $\bar{R}_{n-1}$ (average earning rate) was more positive (Fig. 4.9: LRT=33.283, p<0.001). Additionally, participants were faster to initiate trials as the number of trials completed increased (Fig. 4.9: LRT=93.424, p<0.001).

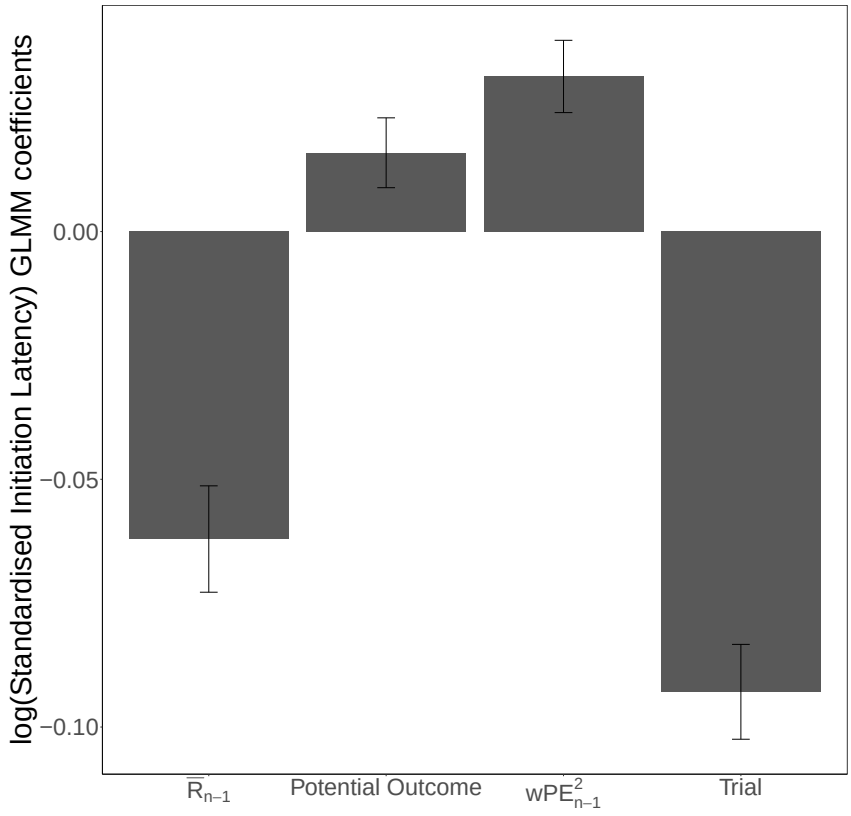


FIGURE 4.9: Standardised GLMM coefficients from the model of the log-transformed latency to initiate each trial. Error bars represent one standard error

4.4 Discussion

In this study we aimed to develop a human judgement bias task which more closely mirrored non-human animal versions of the task. To achieve this, we modified a human judgement bias task that we had previously used (Chapter 3), replacing monetary gain with apple juice, and monetary loss with salty tea. By using a near identical task, and the same computational approach to data analysis, we could assess whether and how replacing secondary with primary reinforcers influenced the relationship between reward and punisher experience, affect, and decision-making.

4.4.1 How does reward and punisher experience influence decision-making?

Consistent with the results of the monetary version of the task (Chapter 3), recent surprise (i.e. the squared weighted prediction error) was a key determinant of decision-making. When recent outcomes were more surprising, participants overall had a reduced reward sensitivity leading to greater risk-aversion. Thus, although the weighted prediction error was found to operate via reward as opposed to loss sensitivity, the direction of the effect was identical between the different versions of the task. In both

cases, this may reflect an attempt to resolve local uncertainty; the higher weighted squared prediction error reflects that the individual's predictions about the environment are unreliable, and so to increase certainty they opt for the safe option which has a known outcome.

The extent to which outcomes were surprising also modulated the participants' belief that the trial would be rewarded. However, this effect had no clear overall direction which likely indicates that the direction of the effect differed between individuals. This perhaps reflects that both the safe and risky actions could ultimately reduce future uncertainty, the latter being over a longer timescale. Participants can reduce uncertainty most straightforwardly by choosing the safe (go) response, since there is no uncertainty about the null outcome that will result. However, choosing the risky (stay) response is necessary to learn and reduce uncertainty about the contingency between the stimuli and outcomes. However, it is not clear why the effect of surprise on reward sensitivity had a clear direction, while the effect of surprise on prior belief was variable between participants. Speculatively, this might suggest that an altered prior belief and reward sensitivity fulfil different (although unknown) functions in terms of optimising decision-making in accordance with environmental conditions. This would be an interesting avenue for future research.

4.4.2 How does reward and punisher experience influence reported affect?

The study revealed that a higher average earning rate was associated with more positive affective valence. This finding further supports the relationship between the average earning rate and affective valence identified using the monetary task (Chapter 3) and extends it to the context of primary reinforcers, thus confirming that absolute reward and punisher experience does modulate momentary subjective affective valence.

This study also raises the issue that, in humans at least, a change in affect does not necessitate shifts in risk-aversion, while changes in risk aversion do not necessarily correspond to an altered affective state.

Yet, in contrast to the monetary task (Chapter 3), we found no evidence that the squared weighted prediction error influenced arousal. This could reflect that the volumes of juice and salty tea delivered were very small and consequently could be less arousing generally than monetary rewards or losses.

4.4.3 How do parameters characterising decision-making relate to affect?

Interestingly, more negative ratings of affective valence were associated with weaker surprise-dependent modulation of reward sensitivity. This is in accordance with the results of the monetary task (Chapter 3) in which decision-making of individuals reporting more negative affective valence were less influenced (albeit via loss aversion)

by the squared weighted prediction error. Hence, this firstly provides further evidence for a potential link between prediction error blunting and affective state, which has been demonstrated neurophysiologically (Gradin et al., 2011; Kumar et al., 2008; Steele et al., 2004). Secondly, it might also provide further support for the concept of cautious optimism which suggests that happier individuals may be more risk-averse in their actual decision-making despite a stronger belief that an outcome will be favourable (Isen et al., 1988; Isen and Patrick, 1983; Nygren et al., 1996). Similarly, in line with cautious optimism, participants who reported more negative affective valence were more likely to have a stronger belief that the 'stay' response would result in salty tea in periods of greater uncertainty. This result may be explained by the finding that poorer mood states can lead to the immediate dismissal of actions that could lead to large punishers, even if they would be beneficial in the longer term (Huys et al., 2012). Here, learning from the outcome of the risky response may ultimately improve performance and increase overall juice intake, but the increased belief that the trial would lead to delivery of salty tea may have increased risk-aversion and decreased willingness to make the risky but potentially beneficial action in periods of uncertainty. It might also reflect that there is an interaction between longer-term mood and environmental conditions, such that individuals in more negative affective states have a greater expectation of punishers (or reduced expectation of rewards) but that this only arises, or is particularly apparent, in periods of uncertainty.

Participants who reported more negative affect were more sensitive to rewards. Although one might typically expect more negative (particularly clinical) affective states to be associated with anhedonia and so a reduced reward value, negative affect has been associated with increased reward value on numerous occasions (Neville et al., 2017; Simmons et al., 2016; Spruijt et al., 2001; Van der Harst et al., 2003). This could reflect a form of mood-repair (Sanchez et al., 2014) whereby individuals seek rewards to enhance their negative mood, which could potentially arise in less severe negative states as opposed to more severe negative states in which anhedonia is observed. It might also reflect cautious optimism; that individuals in more positively valenced affective states are more risk-averse (resulting from reduced reward sensitivity) as a means to maintain their positive affect. This potential explanation has important implications for the judgement bias task as a measure of welfare as it suggests that risk-seeking behaviour may be associated with more negative affective states (although potentially only mild as opposed to severe negative states), and so caution should be taken when interpreting judgement bias results. However, it might generally be the case that negative affect is also associated with a decreased expectation of reward which more than offsets the increased reward sensitivity to produce risk aversion. This has been previously observed (Iigaya et al., 2016; Nygren et al., 1996) although we did not find an affect-dependent prior belief here. Alternatively, an increased reward sensitivity in participants reporting more negative affect could more simply reflect negative affect which resulted from hunger or thirst hence increasing the value of the juice.

4.4.4 The influence of reward and punisher experience on trial initiation

Corroborating the findings of the monetary task (Chapter 3), as well as other studies (e.g. Griffiths and Beierholm, 2017; Guitart-Masip et al., 2011), a higher average earning rate was associated with faster trial initiation. This confirms that participants do attempt to optimise vigour (Niv et al., 2007) in line with reward and punisher experience.

Despite a lack of evidence that surprise (squared weighted prediction error) influenced subjective arousal in this study, it did significantly influence vigour. More specifically, participants were slower to initiate trials when there was greater uncertainty. Previous research has suggested that humans react more slowly following a surprising outcome (Bestmann et al., 2008; Browning et al., 2015) which potentially reflects an attempt to process and learn from the surprising outcomes to reduce subsequent prediction errors and optimise behaviour. However, this contrasts with the results of the monetary task (Chapter 3) in which the weighted prediction error provided a better explanation of variation in vigour than the squared weighted prediction error. This suggests that the main difference between the tasks is in the response to negative prediction errors, with individuals slowing following negative prediction errors in this task but decreasing their latency following negative prediction errors in the monetary task. One potential explanation for this is that losses may fundamentally differ from punishers; it might be worthwhile for a participant to initiate trials with greater vigour to more quickly recoup losses, but as punishers cannot be undone, a better solution in the case of unexpected salty tea or unexpectedly low volumes of juice may be to learn from experience and hence decrease vigour.

4.4.5 The influence of immediate rewards and punishers on behaviour

Across both the primary and secondary (Chapter 3) reinforcer tasks, participants were slower to initiate trials but tended to report greater arousal when the offered volume of juice was greater; perhaps indicating the participants were more alert and took time to prepare on trials with higher stakes. Participants also reported more positive affective valence and were more likely to make the 'stay' response when the potential outcome was higher across both tasks. This indicates that participants were attending to the task, found greater potential juice to be more rewarding, and adjusted their behaviour rationally according to the expected value of the trial, which provides good validation that the task works as anticipated. In future, a within-subject design may help to better understand differences in the influence of primary and secondary reinforcers on behaviour and to determine if there are reliable differences in effect size, allowing us to comment on whether primary rewards and punishers might exert a greater or weaker impact on behaviour than secondary rewards and punishers. Similarly, in accordance with the monetary task (Chapter 3), the previous outcome was a better predictor of

reported valence than the prediction error. Further research could be conducted to see whether this result is generalisable, or whether there are circumstances under which the prediction error better accounts for variation in affective valence.

As observed in the monetary task (Chapter 3), we found a strong effect of the number of trials completed on affective state and behaviour. Participants were reported lower arousal, more negative affective valence, and were faster to initiate trials, and more risk-averse as the task progressed which perhaps suggests that the task induced boredom. This is potentially an issue for the backtranslation of the task given that cognitive tasks have been suggested to be enriching to animals (Krakenberg and Melotti, pers. comm.)

4.5 Conclusions (see Fig. 4.10)

In summary, we have developed a human judgement bias task which uses primary instead of secondary rewards and punishers that more closely resembles the automated task designed for rats (Jones et al., 2018). The results of this study are overall in agreement with that of the human monetary task (Chapter 3) and demonstrate the importance of local uncertainty in decision-making, and how this may relate to affective valence (i.e. affect may play a modulating role), and further supports the average earning rate as an important determinant of affect. There were differences between the human studies in how unexpected outcomes influenced the trial initiation latency which perhaps reflects differences between losses and punishers. The study highlights several issues that would benefit from further research, such as why the average earning rate influences affective valence but not decision-making, and an investigation of the factors which determine the direction in which affect might influence reward sensitivity (e.g. severity of negative affect/timing).

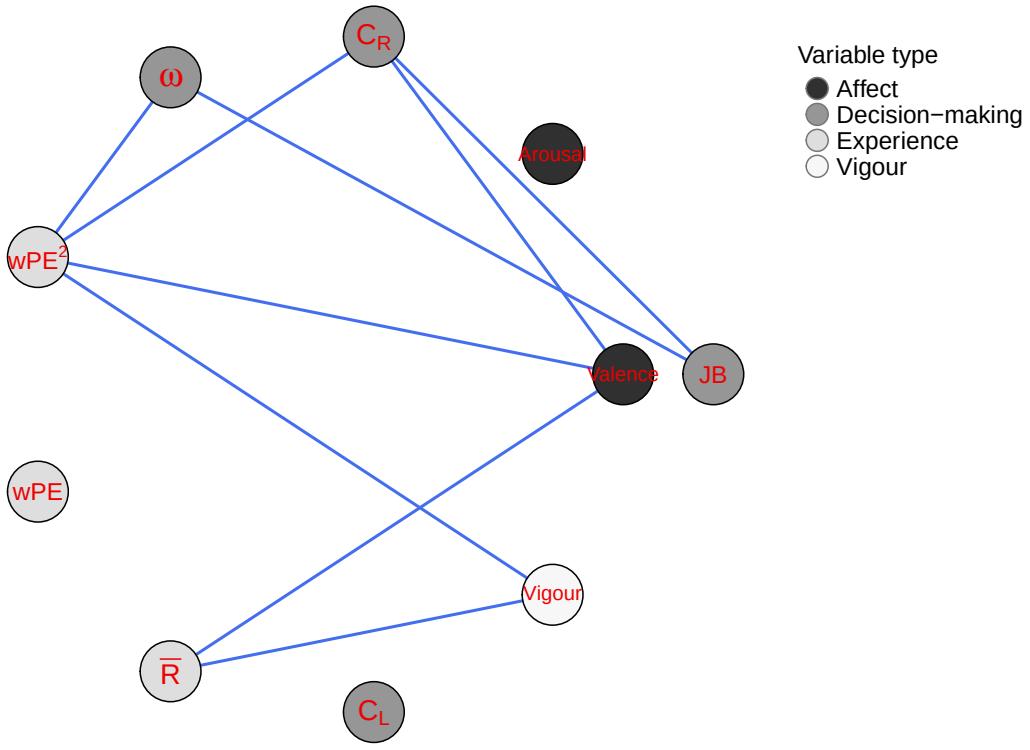


FIGURE 4.10: Diagrammatic summary of the results of Chapter 4: the nodes represent variables relating to affect, decision-making (where C_L denotes loss/punisher sensitivity, C_R denotes reward sensitivity, ω denotes the prior belief about outcomes, and JB denotes judgement bias), experience (where wPE denotes the weighted reward prediction error, $w\text{PE}^2$ denotes the squared weighted reward prediction error, and $\bar{R}$ denotes the average earning/reward rate), and vigour, the lines between nodes represent observed associations between those variables.

Chapter 5

How does pre-test reward and punisher experience influence rat judgement bias and vigour?

Chapter summary: Affect is thought to guide adaptive decision-making. However, it is not clear what information affect encodes about the environment, the situation of an individual within this environment, and how affect influences decision-making. To investigate this, we created environments by presenting either many or few rewards or punishers of different modalities to rats, and then administered two decision-making tasks to quantify the effects of these different environments on choice and motivation. The rewards (sucrose and 50kHz ultrasonic vocalisations) and punishers (air-puffs and 22kHz ultrasonic vocalisations) were designed to induce temporary shifts in affective state. The tasks were a progressive ratio lever pressing task and a judgement bias task. We found that rats pressed a lever to obtain sucrose more frequently in the test sessions following delivery of punishers than rewards. Judgement bias data were analysed in both model-agnostic and model-dependent manners. In contrast to our initial hypothesis, rats who received many sucrose pellets prior to testing were more risk-averse in the judgement bias task than rats who received many air-puffs. The computational analysis revealed that this resulted from a higher weighting of punishers relative to rewards following pre-test sucrose compared to pre-test air-puff experience. While we can conclude that reward and punisher experience alter rat decision-making and vigour, the specific affective states that might arise from such experiences remain unclear.

5.1 Introduction

Certain features of negatively and positively valenced affective states are phylogenetically widespread which suggests that affect confers an evolutionary advantage (Anderson and Adolphs, 2014; Bateson et al., 2011; Darwin, 1872; Mendl et al., 2010; Nettle and Bateson, 2012). A prominent hypothesis is that affect, having both transient and longer-lasting components, provides information about the state of the environment and an individual's situation within it to allow adaptive decision-making (Bach and Dayan, 2017; Mendl et al., 2010; Nettle and Bateson, 2012; Trimmer et al., 2013). More specifically, affect is thought to encode information about reward and punisher experience to optimise future reward acquisition and punisher avoidance (Bach and Dayan, 2017; Mendl et al., 2010; Nettle and Bateson, 2012; Trimmer et al., 2013). This functional framework for affect suggests that decision-making should depend on reward and punisher experience. More specifically, rewards and punishers, regardless of modality, should alter decision-making, with the extent to which they do so being dependent on the prevalence of rewarding and punishing experiences. In general, rewards should be perceived as more likely in environments in which they are frequent, and punishers should be considered more likely in environments in which they are more prevalent. However, to our knowledge, this has yet to be explicitly tested.

A better understanding of the functional significance of affect is crucial to its investigation, particularly in non-human animals. Judgement bias testing is a prominent method for investigating the relationship between affect and decision-making that has been used in numerous species, including rats, starlings, sheep and humans (Doyle et al., 2010; Harding et al., 2004; Iigaya et al., 2016; Matheson et al., 2008). Judgement bias has become a widely used tool for the investigation of affect and welfare in a range of non-human animals (Baciadonna and McElligott, 2015; Mendl et al., 2009; Mendl and Paul, 2004; Yeates and Main, 2008). One issue that complicates the interpretation of judgement bias is a poor understanding of the cognitive mechanisms that contribute to affect-induced changes in judgement bias. Although apparent optimism in the judgement bias task is often assumed to reflect an increased expectation of reward or a decreased expectation of punisher (Bateson, 2016; Mendl et al., 2009), decision-making within the task involves evaluation of both the likelihood and value of potential outcomes (Glimcher et al., 2009; Loewenstein et al., 2008; Neumann and Morgenstern, 1944). Thus, an altered subjective valuation of the reward or punisher could also lead to variation in judgement bias (Mendl et al., 2009). Moreover, there is evidence to suggest that affect might alter both these aspects of cognition (Mendl et al., 2009). Examining these possibilities would aid a better understanding of judgement bias as a measure of affect and welfare and potentially shed light on contradictory findings in the judgement bias. In particular, it could elucidate the circumstances that lead individuals in putatively negative affective states to be more risk-seeking.

The aim of this study was therefore to investigate the putative adaptive function of affect, namely as a Bayesian prior over future outcomes, and to gain a better

understanding of judgement bias as a measure of affect and welfare. Our central hypotheses were that reward and punisher experience would determine the affective (specifically 'emotional') state of an individual according to Roll's (2013) definition of affect (i.e. rewards induce positive affect; and punishers induce negative affect) and that the resulting affective state would influence decision-making and vigour. Given this, we aimed to address six specific questions about the relationship between reward and punisher experience, affect, and decision-making: (1) are individuals more risk-seeking in the judgement bias task following experience of rewards compared to punishers; (2) is this effect dependent on reward and punisher prevalence; (3) does this effect also depend on the extent to which the modality of the rewards and punishers experienced matches that of the potential outcome of the decision; (4) does this risk-seeking behaviour arise from modulation of the likelihood or the value of rewards and punishers, (5) what is the relationship between reward and punisher experience and vigour; and (6) as self-determined inter-trial intervals gave us an opportunity to examine vigour across the test session, does within-test reward and punisher experience alter decision-making and vigour?

To answer these questions, we administered two different sucrose delivering decision-making tasks to animals who had been exposed to one of eight different environments, and used two structurally different analysis methods to examine the results. The tasks were a standard progressive ratio lever pressing task (Hodos, 1961) and an automated judgement bias task in which trials were self-initiated (Jones et al., 2018). The environments crossed valence (reward versus punisher), modality (auditory versus sucrose or air puff unconditioned stimuli) and prevalence (low versus high). We employed a model-agnostic analysis to summarize the raw data; and fit a Bayesian decision-theoretic model (Iigaya et al., 2016) which allowed us to estimate parameters relating to the decision-making process, namely reward valuation and probability estimation, as well as modelling the duration of the self-determined inter-trial intervals to assess how measures of pre-test and within-test reward and punisher experience altered vigour. The progressive ratio lever pressing task provided a behavioural measure of reward valuation to complement the measure of reward and punisher valuation obtained through the model-dependent analysis of the judgement bias data.

We hypothesised that rats would be more risk-seeking immediately following reward compared to punisher experience, regardless of modality (i.e. there would be a significant effect of manipulation valence but not a significant interaction between manipulation valence and manipulation modality on judgement bias, Fig. 5.1). Additionally, we hypothesised that the prevalence of reward and punisher experience would modulate the extent to which the rat exhibited risk-seeking behaviour; rats would be most risk-seeking following experience of rewards at a high prevalence, and least risk-seeking following experience of punishers at a high prevalence (i.e a there would be a significant interaction between manipulation valence and manipulation prevalence on judgement bias, Fig. 5.1). We expected that the model-dependent analysis would reveal that relatively risk-seeking behaviour would arise from a bias towards executing

the risky response. Given the conflicting evidence in the literature about the relationship between experience, affect, and reward and punisher valutation (see Chapter 1), we were agnostic as to whether and how the rats' valuation of rewards/punishers would depend on reward and punisher experience prior to testing. However, we expected that the results of the progressive ratio lever pressing task would be consistent with the results of the model-dependent analysis; rats that had a greater value of sucrose compared to air-puffs according to the model parameter characterising the relative sensitivity of punishers to rewards, would also make a greater number of lever presses to obtain sucrose. With regards to the inter-trial interval data, in line with Beierholm et al. (2013); Niv et al. (2007), we hypothesised that rats would be faster to initiate a trial when the average reward rate was higher. Also, we expected that if rats became satiated as a result of an increasing intake of sucrose, this be apparent as an effect of trial on the inter-trial interval (i.e. increased latencies as the number of trials completed increased).

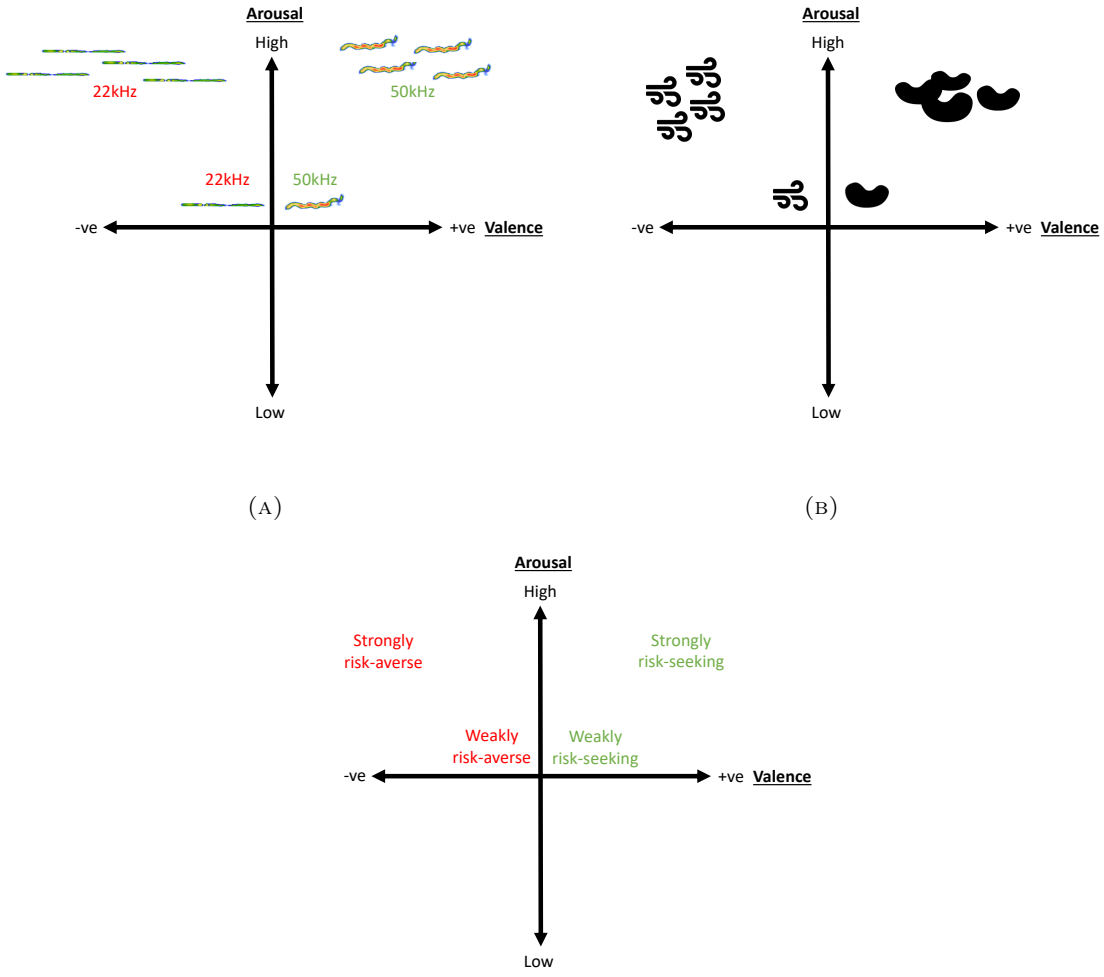


FIGURE 5.1: Diagrammatic summary of hypotheses; experiencing rewards and punishers non-specific (A) and specific (B) to the task would influence affect; these predicted affective states would be associated with changes in risk-aversion (C).

5.2 Methods

5.2.1 Subjects

Subjects were 16 male Lister Hooded rats (Charles River, Margate, UK). The rats were housed in stable pairs in cages measuring 560x340x190mm and kept under a 12-hour reverse light cycle, with lights off at 7am and on at 7pm. Two cardboard tubes and an aspen block were provided as enrichment and subjects had ad libitum access to food (LabDiet) and water. This research was conducted under University Investigation Number UB/16/004 and adhered to ASAB/ABS guidelines for the use of animals in research. Subjects were rehomed as pet rats on completion of the experiment.

5.2.2 Procedure

Subjects first completed training and testing on the progressive ratio lever pressing task, before being trained and tested on the judgement bias task. Prior to each test session, rats received rewards and punishers at a high or low prevalence. Once rats had met the criterion to progress to testing for each task, test sessions occurred twice per week, on a Tuesday and Thursday, with additional training sessions on Monday, Wednesday, and Friday. Rats were not food or water restricted prior to training or testing.

5.2.3 Apparatus (see Fig. 5.2)

Four identical shuttle boxes measuring 508×254×305mm placed in acoustic isolation chambers were used for both the judgement bias and progressive ratio lever pressing task. A wall divided each box into two sections which each measured 254×254×305mm. Training and testing of each task occurred in only one of these sections. During the progressive ratio lever pressing task a retractable lever was attached to the centre of the end wall of each operant box and a plastic pot was attached to a pellet feeder was placed adjacent to this lever. For the judgement bias task, the retractable lever was replaced with a pellet/air-puff delivery trough, the pot was removed, and a speaker was placed on top of each operant box facing down into the operant box. The pre-test treatment took place in one of two identical operant boxes, both measuring 305×178×355mm in separate isolation chambers in separate rooms adjacent to the testing room. A plastic pot was attached to the end wall in both boxes. This pot was attached to a pellet feeder in one box, and an ultrasonic speaker was placed in the chamber containing the other box. The operant equipment was manufactured by Coulbourn Instruments (Allentown, PA, USA), and operated by Graphic State (v4) software. The ultrasonic vocalisations (USVs) were broadcast using Avisoft software and ultrasound gate hardware (Avisoft Bioacoustics, Berlin, Germany). The sucrose pellets used throughout the experiment were Bioserv (Frenchtown, NJ, USA) Dustless Precision Pellets (45 mg).

FIGURE 5.2: The shuttle box within a sound isolation chamber used for the judgement bias task; there is a pellet dispenser connected to the trough on the end wall.

5.2.4 Progressive ratio lever pressing task

To train rats to press a lever to obtain sucrose, each lever press resulted in the lever retracting for one second and a sucrose pellet being delivered into the pot. An autoshaping procedure ran simultaneously, such that every 60 seconds the lever would retract and a sucrose pellet would be delivered, even if the lever had not been pressed, after which the lever became re-available (Brown and Jenkins, 1968; Parker et al., 2014). Initially, the duration of the training sessions was 15 minutes. Once subjects had completed at least three 15-minute training sessions and had made at least 15 lever presses in a session, the duration of training sessions was increased to 30 minutes. Rats progressed to testing following completion of two 30-minute training sessions in which at least 30 lever presses were made. During testing, the number of presses required before the sucrose pellet was delivered increased in increments of five lever presses starting at one lever press. The test session ended after 30 minutes had elapsed.

5.2.5 Judgement bias task

The judgement bias task followed the methodology of Jones et al. (2018); see Fig. 5.3. Judgement bias task trials were self-initiated by the rat inserting his snout into the food delivery trough which led to a tone being played. Rats were trained to keep their snout in the trough for two seconds when the positive tone played to receive a sucrose pellet, and remove it when the negative tone played to avoid an air-puff. Removal of the snout before two seconds had elapsed resulted in no sucrose or air-puff being

delivered regardless of the tone presented. The reference tones had a frequency of 2kHz or 8kHz, and the positive tone was 2kHz for half of the rats and 8kHz for the remaining half. In test sessions, three probe tones (2.8kHz, 4kHz, 5.6kHz) were presented on six trials each in addition to the reference tones on 21 trials each. The order of tone presentation was randomised and the ambiguous tones were non-reinforced.

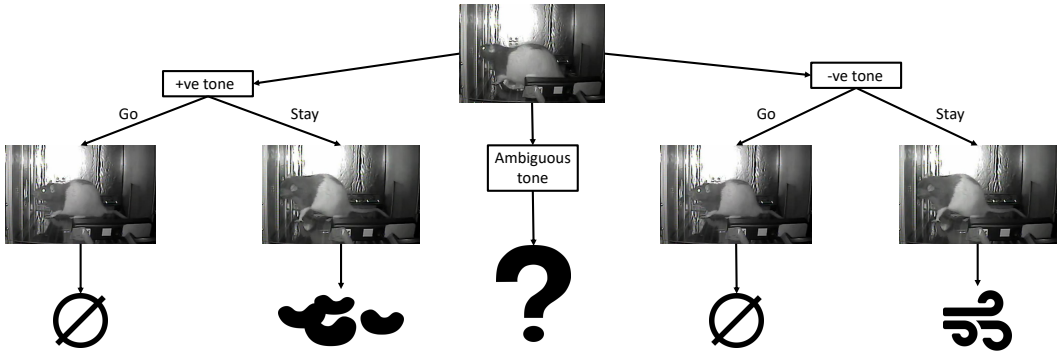
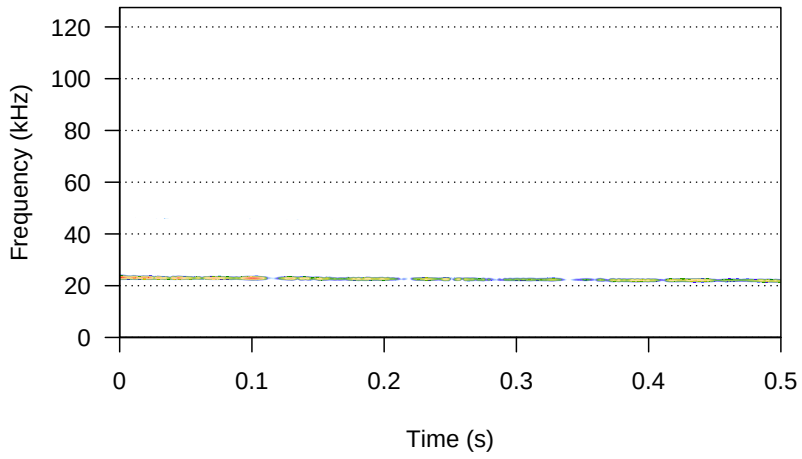


FIGURE 5.3: An example of judgement bias test, using the task described by Jones et al. (2018): following trial initiation, the rat is presented with a tone which is either +ve, -ve, or ambiguous. If they execute the risky response ('stay') when the +ve tone is presented they will be rewarded with food, but executing this response when the -ve tone is presented will result in a punishing air-puff. The safe response ('go') results in no reward or punisher. During testing, the rat is presented with several ambiguous tones and must decide whether to execute the 'stay' or 'go' response

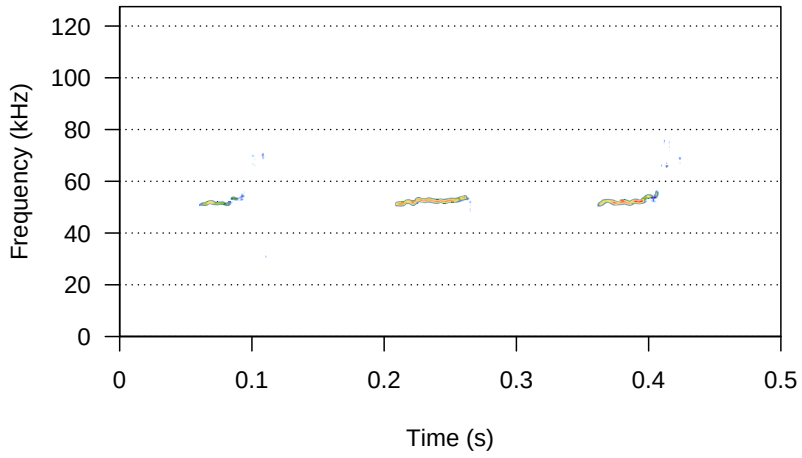
5.2.6 Affect manipulation

Reward or punisher experience was manipulated in the 15 minutes prior to each testing session. We adopt Rolls' (2013) definition of rewards as anything that an animal will work for, and a punisher as anything that an animal will attempt to escape from or avoid. The rewards delivered were one sucrose pellet or playback of a 50kHz USV lasting eight seconds, and the punishers delivered were an air-puff directed towards the rat that was stopped as soon as the rat moved away or playback of a 22kHz USV of eight second duration (Fig. 5.4). Rats will work to hear playback of 50kHz USVs, which are emitted by rats in situations considered to be affectively positive, and will work to avoid playback of 22kHz USVs which are emitted by rats during unpleasant experiences (Burgdorf et al., 2008). Similarly, rats will work to obtain sucrose pellets, and will actively escape air-puffs (Cimadevilla et al., 2001; Engelmann et al., 1996; Sclafani and Ackroff, 2003). Each reward or punisher was delivered at either a high or low prevalence. The reward or punisher was delivered at 15 randomly selected times in the high prevalence condition, and once at a randomly selected point during the low prevalence condition. The experimental design was within-subject and therefore each rat experienced eight manipulations of reward or punisher experience per task over four weeks. Rats experienced the same reward or punisher at different prevalences within the same week, and the order of rewards or punishers experienced was unique

for each rat on each task and was assigned randomly. Each rat experienced the high prevalence condition on the first test day of the week for half the test weeks which were selected at random for each rat. Except for the air-puff which was administered by the experimenter using an air-cannister (Fellowes, Doncaster, UK: Air Duster 300ml), delivery of the rewards and punishers was automated. The 50kHz USV playback and 22kHz playback were recorded from rats unknown to the experimental subjects. The 50kHz vocalisations were induced by placing the rats back into their home cage following a brief period in a neutral holding cage. The 22kHz USVs were induced by placing the rats in an open arena.



(A)



(B)

FIGURE 5.4: Spectrogram of excerpt from USV recordings used for the affect manipulation: (A) the 22 kHz playback; (B) the 50kHz playback.

5.2.7 Data analysis

Two rats did not complete judgement bias testing due to poor performance in training. Data analysis comprised two approaches: a model-agnostic and model-dependent approach. Choice data from the judgement bias task were analysed using both approaches which assessed whether there were treatment differences in judgement bias (model-agnostic analysis) and explored the potential cognitive processes underlying these treatment differences (model-dependent analysis). Data from the progressive ratio lever pressing task were only analysed using a model-agnostic approach to assess whether there were treatment differences in the number of lever presses made in a session, which could be compared to the parameter characterising sensitivity to rewards and punishers from the model-dependent analysis. These data were unsuitable for analysis using a model-dependent approach as each test session provided only one data point (i.e. number of lever presses). Self-determined interval data were analysed solely using a model-agnostic approach which assessed whether overall treatment differences existed in self-determined inter-trial intervals, examined which variables relating to reward and punisher experience might contribute to variation in inter-trial intervals, and examined whether the influence of these variables depended on treatment.

Model-agnostic analysis

The response variables analysed were: the total number of lever presses in the progressive ratio lever pressing task, and the decision to 'stay' (risky) or 'go' (safe) in the judgement bias task. These data were analysed using generalised linear mixed models (GLMM) implemented in the lme4 (Bates et al., 2015) and nlme (Pinheiro et al., 2018) packages in R (R Core Team, 2017). Where models assumed a Gaussian error function, the residuals of each model were visually inspected to verify that the model assumptions of normality of error and homogeneity of variance were met. Following log-transformation of the total number of lever presses, these assumptions were not violated in any model. Likelihood-ratio tests (LRT) were used to assess whether the difference in model deviance was significant when a parameter was removed from the model.

These GLMMs included a random effect of individual and session, with session nested within individual. All models included the following fixed effects: manipulation valence (reward or punisher), manipulation modality (USV or task specific), manipulation prevalence (high or low), and interactions between manipulation valence, modality, and prevalence. Additional fixed effect terms included in the model of decision were: a variable which coded whether the 2kHz or 8kHz tone was used as the rewarded test stimulus, number of trials completed prior to the decision, the tone presented, and the prediction error on the previous trial (as determined by the model-dependent analysis of the self-determined inter-trial intervals), the previous outcome,

and interactions between stimulus and valence, and valence, prevalence, and modality. To investigate significant and marginally non-significant interaction terms in the model, post-hoc comparisons were conducted using a simultaneous pairwise Tukey procedure for general linear hypotheses within the R package multcomp (Hothorn et al., 2008).

Model-dependent analysis: judgement bias

We used a Bayesian-decision theoretic model to characterise decision-making in the judgement bias task (Iigaya et al., 2016; Whiteley and Sahani, 2008). On each trial of the task, subjects are presented with a tone (s) which we consider to take values between -2 and 2, where -2 represents the punished reference tone, 2 represents the rewarded reference tone, and 0 is equidistant from the reference tones and is thus entirely ambiguous. The subject's perception of the tone, x, will inform their estimation of the probability that the tone signals the delivery of a punisher, assuming that $s < 0$ is associated with a punisher, and $s > 0$ with a reward:

$$P(\text{pun}|x) = \int_{-\infty}^{0} ds\ P(s|x) \tag{5.1}$$

The expected value of making the 'stay' response depends on both the probability that the tone presented (s) signals a test punisher, given the subject's perception of the tone (x) and the subject's subjective value of the test sucrose and air-puff, c^+ and c^- respectively:

$$E_{p/r}(x) = c^+(1 - P(\text{pun}|x)) - c^- P(\text{pun}|x) \tag{5.2}$$

The expected reward for making the 'go' response is 0. Thus, assuming that the posterior probability of s given x follows a Gaussian distribution with mean x and variance σ^2, i.e. $P(s|x) \sim \mathcal{N}(s : x, \sigma^2)$, it is optimal for a subject to make the 'stay' response under the following condition:

$$c^+(1 - P(\text{pun}|x)) - c^- P(\text{pun}|x) > 0 \tag{5.3}$$

$$c^+(1 - \phi_\sigma(-x)) - c^- \phi_\sigma(-x) > 0 \tag{5.4}$$

$$x > -\phi_\sigma^{-1}(\alpha) \tag{5.5}$$

where:

$$\alpha = \frac{1}{1 + C_{p/r}} \tag{5.6}$$

and $C_{p/r} = \frac{c^-}{c^+}$ is the differential sensitivity of the air-puff to sucrose (Fig. 5.5). The probability that a subject opts to 'stay' given the true tone is therefore:

$$P_{stay} = P(E_{p/r}(x) > 0|s) \tag{5.7}$$

$$= \phi_\sigma(s + \phi_\sigma^{-1}(\alpha)) \tag{5.8}$$

We include two additional parameters: a lapse rate λ to account for non-perceptual errors, and a bias term δ to allow biases towards or away from 'stay' response which thus provides a representation of risk-aversion (Fig. 5.5; Table 5.1). The probability of a subject making the 'stay' response on any given trial is given as:

$$P_{stay} = (1 - \lambda)\phi_\sigma(s + \phi_\sigma^{-1}(\alpha) + \delta) + \frac{\lambda}{2} \tag{5.9}$$

A relatively risk-seeking judgement bias can therefore arise from either values of $C_{p/r}$ that are less than 1 or values of δ greater than 0.

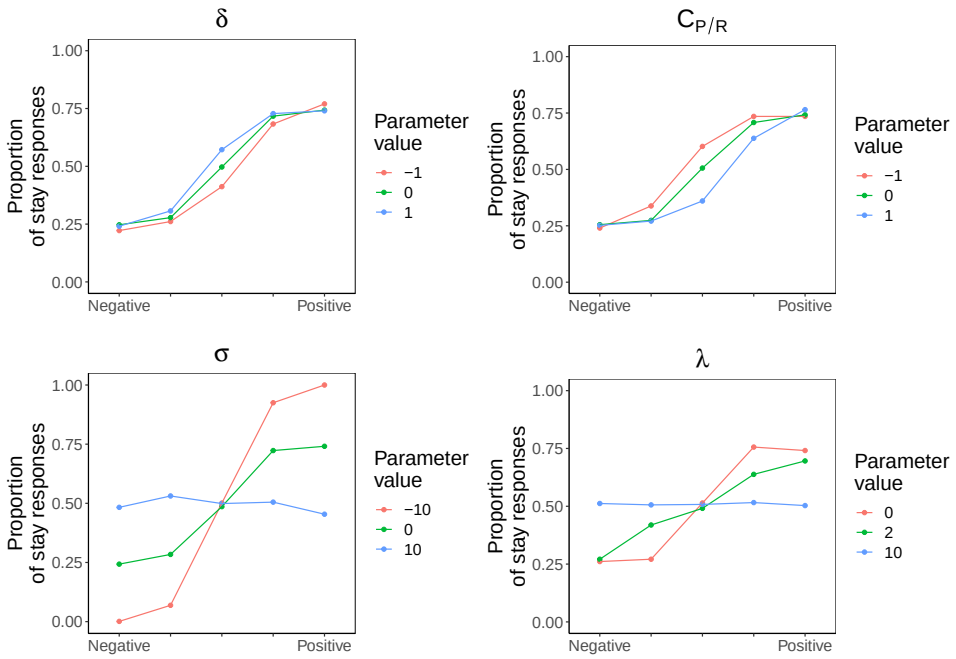


FIGURE 5.5: Data generated using the described model to demonstrate parameter-dependent variation in judgement bias for each of the model parameters

TABLE 5.1: Glossary of judgement bias model parameters

Parameter	**Range**	**Interpretation**
$C_{P/R}$	$[0, \infty]$	Relative scaling of punishers to rewards: higher values reflect that avoiding the air-puff is of greater value than obtaining a sucrose pellet
δ	$[-\infty, \infty]$	Bias: higher values reflect a bias towards executing the 'stay' response
λ	$[0, 1]$	Lapse rate: higher values a greater likelihood of executing the action with the lowest value (i.e. the 'wrong' action).
σ	$[0, \infty]$	Slope parameter: higher values reflect a poorer ability to detect the true direction of the stimulus.

Model-dependent analysis: self-determined inter-trial intervals

The log-transformed self-determined inter-trial interval data was modelled using linear regression with a constant term k and the following regressors: weighting for the previous outcome ($\beta_{R_{t-1}}$), where the previous outcome (R_{t-1}) was either 1 (sucrose), 0 (nothing), or $-C_{p/r}$ (air-puff scaled by relative sensitivity of losses to rewards), weighting for the prediction error ($\beta_{\mathrm{PE}_{t-1}}$), weighting for the squared prediction error ($\beta_{\mathrm{PE}^2_{t-1}}$), and a weighting for the number of trials completed (β_t).

The prediction error on each trial was calculated as the difference between R_t and the average of the potential outcomes:

$$\mathrm{PE}_{(t)} = R_t - \left(\frac{1}{3} - \frac{C_{p/r}}{3} + \frac{0}{3}\right) \tag{5.10}$$

This was found to provide a better fit than a prediction error based on the prediction following stimulus presentation (ΔAIC=510.050, comparing all models fitted calculating the prediction error as above the same set of models that instead calculated the prediction error as below):

$$\mathrm{PE}_{(t)} = R_t - ((1 - P(\mathrm{pun}|x_t) - C_{p/r}(\mathrm{pun}|x_t)) \tag{5.11}$$

where $P(\mathrm{pun}|x_t)$ is sigmoidal and dependent on the ability of the rat to discriminate between the stimuli (σ) and the stimulus presented on the trial (s_t):

$$P(\mathrm{pun}|x_t) = \phi_\sigma(s_t + \phi_\sigma^{-1}(0.5)) \tag{5.12}$$

The linear regression model is hence given as:

$$log(ITI) = k + (\beta_{R_{t-1}} \times R_{t-1}) + (\beta_{\mathrm{PE}_{t-1}} \times \mathrm{PE}_{t-1}) + (\beta_{\mathrm{PE}^2_{t-1}} \times \mathrm{PE}^2_{t-1}) + (\beta_t \times t) + \epsilon \tag{5.13}$$

where ϵ is assumed to be normally distributed with mean zero and variance ς ($\epsilon \sim \mathcal{N}(0, \varsigma)$).

Model Fitting

Model fitting was conducted using a three-level hierarchical Bayesian random effects analysis. We assume that the parameters $\mathbf{h}^i_j$ for each subject $i \in \{1, \ldots, M\}$ across test sessions $j \in \{1, \ldots, N^i\}$ are a random sample from a Gaussian distribution $\mathbf{h}^i_j \sim \mathcal{N}(\boldsymbol{\mu}^i, \boldsymbol{\Sigma})$ parameterised by $\boldsymbol{\mu}^i$, which is itself a random sample from a Gaussian distribution $\boldsymbol{\mu}^i \sim \mathcal{N}(\mathbf{m}, \boldsymbol{\nu})$ parameterized by $\mathbf{m}$ and $\boldsymbol{\nu}$ (with diagonal $\boldsymbol{\nu}$). Here, $\boldsymbol{\Sigma}$ is also diagonal, and is another parameter.

The aim of model-fitting is to identify the parameters which maximize the likelihood of the data, $D = \{D_j^i\}$ for session j of rat i:

$$\boldsymbol{\Sigma}^{ML}, \mathbf{m}^{ML}, \boldsymbol{\nu}^{ML} \approx \text{argmax}_{\boldsymbol{\Sigma},\mathbf{m},\boldsymbol{\nu}}\{P(D|\boldsymbol{\Sigma}, \mathbf{m}, \boldsymbol{\nu})\} \tag{5.14}$$

$$= \text{argmax}_{\boldsymbol{\Sigma},\mathbf{m},\boldsymbol{\nu}}\left\{ \prod_{i=1}^{M}\prod_{j=1}^{N^i} \int_{\boldsymbol{\mu}^i}\int_{h_j^i} P(\boldsymbol{\mu}^i|\mathbf{m},\boldsymbol{\nu}) P(\mathbf{h}_j^i|\boldsymbol{\mu}^i,\boldsymbol{\Sigma}) P(D_j^i|\mathbf{h}_j^i) d\boldsymbol{\mu}^i d\mathbf{h}_j^i \right\} \tag{5.15}$$

We estimated these parameters using an expectation-maximisation procedure, which involved iterating between the E and M steps described below until convergence (Guitart-Masip et al., 2012). This is a common procedure for the fitting of behavioural data (Guitart-Masip et al., 2012; Iigaya et al., 2016). For the k^{th} iteration of the E-step, we use a Laplacian approximation based on the model parameters that maximized the likelihood of the data given the prior distributions (for convenience, not integrating out $\boldsymbol{\mu}^i$, as would be possible under the current circumstances):

$$\bar{\boldsymbol{\mu}}^i(k), \{\bar{\mathbf{h}}_j^i(k)\} = \text{argmax}_{\boldsymbol{\mu},\{\mathbf{h}_j\}}\left\{ \log P(\boldsymbol{\mu}^i|\mathbf{m}(k),\boldsymbol{\nu}(k)) + \sum_{j=1}^{N^i} \log P(\mathbf{h}_j^i|\boldsymbol{\mu}^i;\boldsymbol{\Sigma}(k)) P(D_j^i|\mathbf{h}_j^i) \right\} \tag{5.16}$$

The inverse Hessian of this maximization leads to M covariance matrices, the i^{th} of which, $\mathcal{M}^i(k)$ has dimensions $(1{+}N^i)h{\times}(1{+}N^i)h$ where h is the number of parameters (with rows and columns organized as $\boldsymbol{\mu}^i, \mathbf{h}_1^i, \mathbf{h}_2^i, \ldots, \mathbf{h}_{N^i}^i$). We write $\mathcal{M}_{xy}^i(k)$ for the $x, y^{\text{th}}, [x, y \in \{0, 1, \ldots N^i\}]$ block of this matrix, with $\mathcal{M}_{00}^i(k)$ being the covariance of $\boldsymbol{\mu}^i(k)$, $\mathcal{M}_{jj}^i(k)$ being the covariance of $\mathbf{h}_j^i(k), [j \in \{1, \ldots, N^i\}]$, and $\mathcal{M}_{0j}^i(k)$ being the covariance between $\boldsymbol{\mu}^i(k)$ and $\mathbf{h}_j^i(k)$, all on iteration k.

The parameters $\boldsymbol{\Sigma}(k)$, $\mathbf{m}(k)$, and $\boldsymbol{\nu}(k)$ are then updated in the M-step as follows:

$$\mathbf{m}(k{+}1) = \frac{1}{M}\sum_{i=1}^{M} \bar{\boldsymbol{\mu}}^i(k) \qquad , \qquad \mathbf{h}^i(k{+}1) = \frac{1}{N^i}\sum_{j=1}^{N^i} \bar{\mathbf{h}}_j^i(k) \tag{5.17}$$

$$\boldsymbol{\nu}(k{+}1) = \text{diag}\left\{ \frac{1}{M}\sum_{i=1}^{M} \left(\bar{\boldsymbol{\mu}}^i(k)[\bar{\boldsymbol{\mu}}^i(k)]^T + \mathcal{M}_{00}^i(k)\right) - \mathbf{m}(k{+}1)[\mathbf{m}(k{+}1)]^T \right\} \tag{5.18}$$

$$\boldsymbol{\Sigma}(k{+}1) = \text{diag}\Bigg\{ \frac{1}{\sum_{i=1}^{M} N^i} \Bigg[\sum_{i=1}^{M} \Bigg(\sum_{j=1}^{N^i} \bar{\mathbf{h}}_j^i(k))[\bar{\mathbf{h}}_j^i(k)]^T - 2\bar{\mathbf{h}}_j^i(k))[\bar{\boldsymbol{\mu}}^i(k)]^T + \bar{\boldsymbol{\mu}}^i(k)[\bar{\boldsymbol{\mu}}^i(k)]^T + \tag{5.19}$$

$$\mathcal{M}_{jj}^i(k) - 2\mathcal{M}_{0j}^i(k) + \mathcal{M}_{00}^i(k) \Bigg) - N^i(\mathbf{h}_j^i(k{+}1) - \bar{\boldsymbol{\mu}}^i(k))[(\mathbf{h}_j^i(k{+}1) - \bar{\boldsymbol{\mu}}^i(k))]^T \Bigg] \tag{5.20}$$

Model-fitting was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bris.ac.uk/acrc/. This involved using one node of the high-performance computer (BlueCrystal Phase

3) to fit each model the data, with multiple models fitted simultaneously, allowing all models to be fitted in under 300 hours. The likelihood function and code for the EM procedure were written using MATLAB. The initial values for all parameters were a number drawn at random between zero and one. Values for parameters constrained between zero and ∞ were exponentiated within the likelihood function, likewise values for parameters constrained between zero and one were transformed using a logistic function within the likelihood function.

An alternative approach to model-fitting would be characterise the full posterior distributions of the parameters. However, even with use of the high-performance computer, this would have taken an an impractical amount of time.

Model comparison

We compared models according to their Akaike Information Criterion (AIC) scores:

$$\text{AIC} = 2N_p - \left(2\sum_{i=1}^{M}\sum_{j=1}^{N^i} logP(D|\mathbf{h}_j^i)\right) \tag{5.21}$$

where N_p is the number of fitted parameters.

Permutation test and condition comparisons

We used a permutation test to assess whether parameter estimates differed from zero, where this was meaningful. This involved calculating the mean parameter estimate and multiplying randomly selected parameter estimates by minus one and recalculating the mean parameter estimate 10000 times. A p-value was obtained by calculating the proportion of resampled means with an absolute difference greater than or equal to the observed mean.

5.3 Results

5.3.1 Statistical analysis: progressive ratio lever pressing data

When rats had experienced punishers prior to the progressive ratio lever pressing task they pressed the lever more than when they had experienced pre-test rewards (mean number of lever presses ± SE: reward: 159±20.4, punishers: 167±14.1, LRT=4.944, p=0.026). The numerical prevalence of pre-test rewards and punishers (LRT<0.001, p=0.993), or modality (LRT=1.927, p=0.165) did not significantly predict the number of lever presses. There was a marginally non-significant interaction between manipulation valence and prevalence (Fig. 5.6: LRT=3.003, p=0.083). While there was a significant difference between rats that experienced high prevalence rewards and high prevalence punishers (mean number of lever presses ± SE: reward: 159.094±32.076, punishers: 173.594±20.164, LRT=-2.913, p=0.013), no significant difference was observed between rats that experienced rewards and punishers at a low prevalence

(LRT=0.421, p=0.967), between rats experiencing rewards at a high or low prevalence (LRT=1.248, p=0.533), or punishers at a high or low prevalence (LRT=-1.247, p=0.534). There was no significant interaction between manipulation valence and modality (LRT= 0.074, p=0.785), or between manipulation valence, modality, and prevalence (LRT= 0.312, p=0.856).

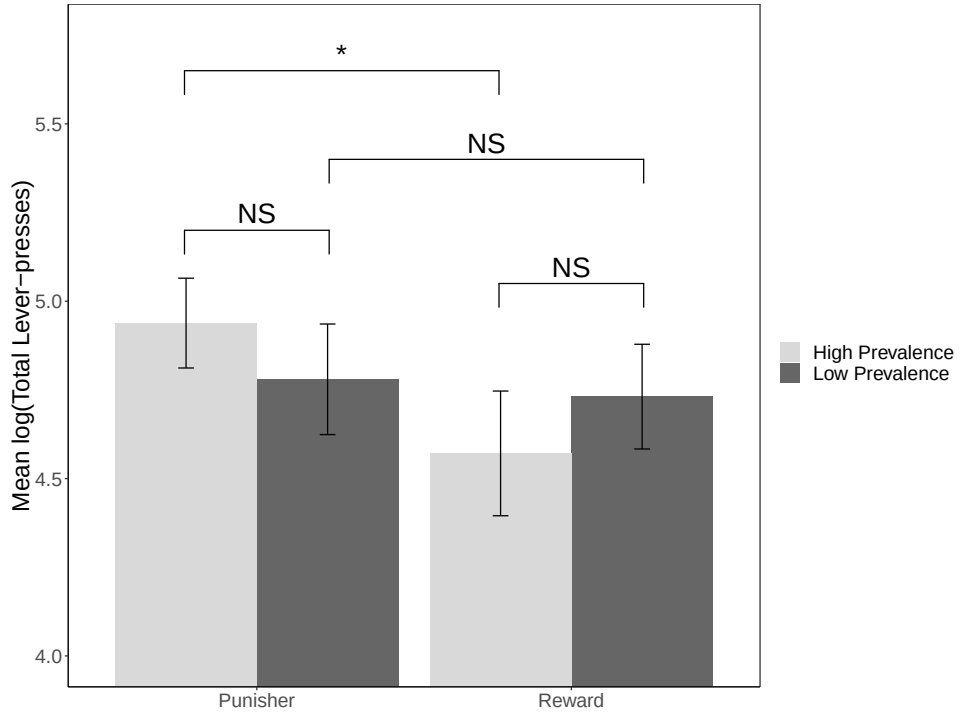


FIGURE 5.6: The effect of manipulation valence and prevalence on the log-transformed number of lever presses collapsed across modality. Error bars represent one standard error

5.3.2 Statistical analysis: judgement bias data

There was a high degree of correlation between the weighted prediction error, squared prediction error, and previous outcome across individuals. To avoid issues arising from multicollinearity, GLMMs containing only one of these variables were compared according to their AIC values and the GLMM with the lowest AIC value was selected for further analysis. The model which included the previous outcome instead of the squared weighted prediction error or weighted prediction error provided a marginally improved fit (ΔAIC=1.744; comparing a saturated model containing the previous outcome and not the squared weighted prediction error or weighted prediction error with the next best saturated model which contained the weighted prediction error but not the previous outcome or squared weighted prediction error) and hence the previous outcome was included as a predictor variable in the GLMM. The stimulus tone was a significant predictor of response, indicating that rats were able to discriminate between the stimuli used in the judgement bias task (LRT=3387.919, p<0.001). Rats were more risk-seeking when the previous outcome

was more aversive (LRT=4.604, p=0.032) and when the 2Khz compared with the 8kHz tone was rewarded (LRT=14.716, p<0.001). Valence (LRT=2.451, p=0.117), modality (LRT=0.092, p=0.762), prevalence (LRT=0.756, p=0.385), and number of trials completed (LRT=2.263, 0.132) were not significant as main effects. However, there was a marginally non-significant interaction between valence and prevalence (Fig. 5.7: LRT=3.467, p=0.063) and a significant interaction between valence and modality (Fig. 5.8: LRT=6.682, 0.010). Post-hoc analyses found a marginally non-significant difference in risk-aversion between rats that experienced highly prevalent prior punishers and highly prevalent pre-test rewards (mean proportion 'stay' responses ± SE: highly prevalent rewards 0.506±0.031 vs. highly prevalent punishers 0.591±0.031, z=-2.302, p=0.074) and no significant difference between rats that experienced pre-test rewards and punishers at a low prevalence (z=-0.331, p=0.984). Additionally, there was no significant difference between high and low prevalence pre-test reward conditions (z=1.994, p=0.150), and no significant difference between high and low prevalence pre-test punisher conditions (z=-0.632, p=0.900). Rats made significantly fewer risk-seeking responses in the pre-test sucrose condition compared to the pre-test air-puff condition (mean proportion 'stay' responses ± SE: sucrose 0.520±0.031 vs. air-puff 0.604±0.031, z=-2.763, p=0.021), but there was no significant difference between the two pre-test USV conditions (z=-0.755, p=0.843). There was no significant difference in judgement bias after experiencing pre-test sucrose compared with 50kHz USVs (z=1.435, p=0.413) and also no significant difference in judgement bias following pre-test air-puffs compared with 22kHz USVs (z=-2.103, p=0.117).

The interaction between valence and stimulus was non-significant (LRT=0.070, p=0.791) as was the interaction between manipulation valence, prevalence, and modality (LRT=0.795, p=0.672).

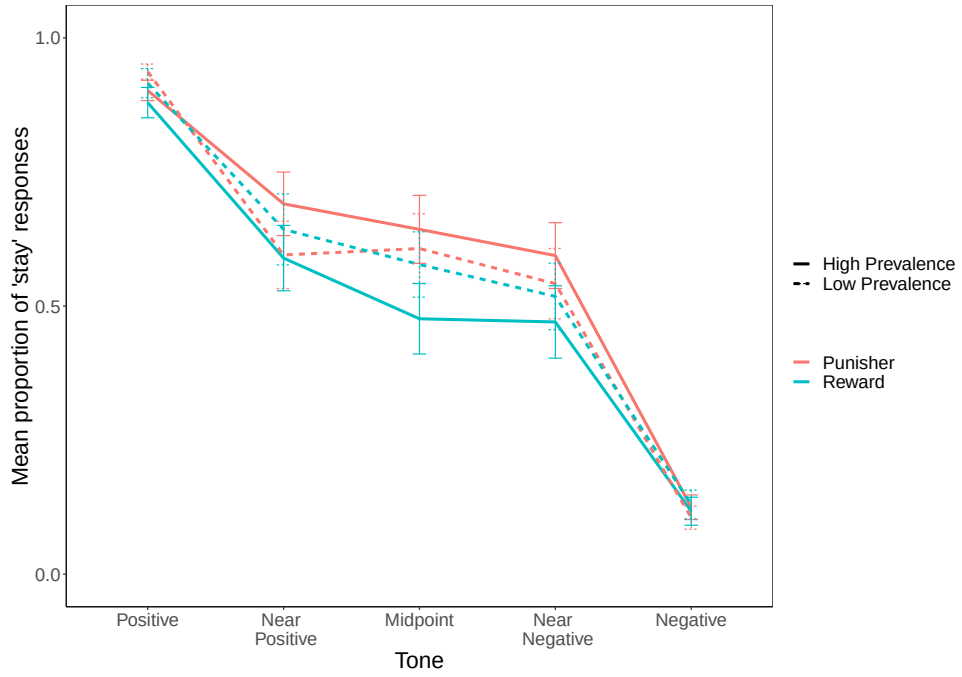


FIGURE 5.7: Mean proportion of risk-seeking responses split by tone presented following experience of high or low prevalence rewards or punishers. Error bars represent one standard error

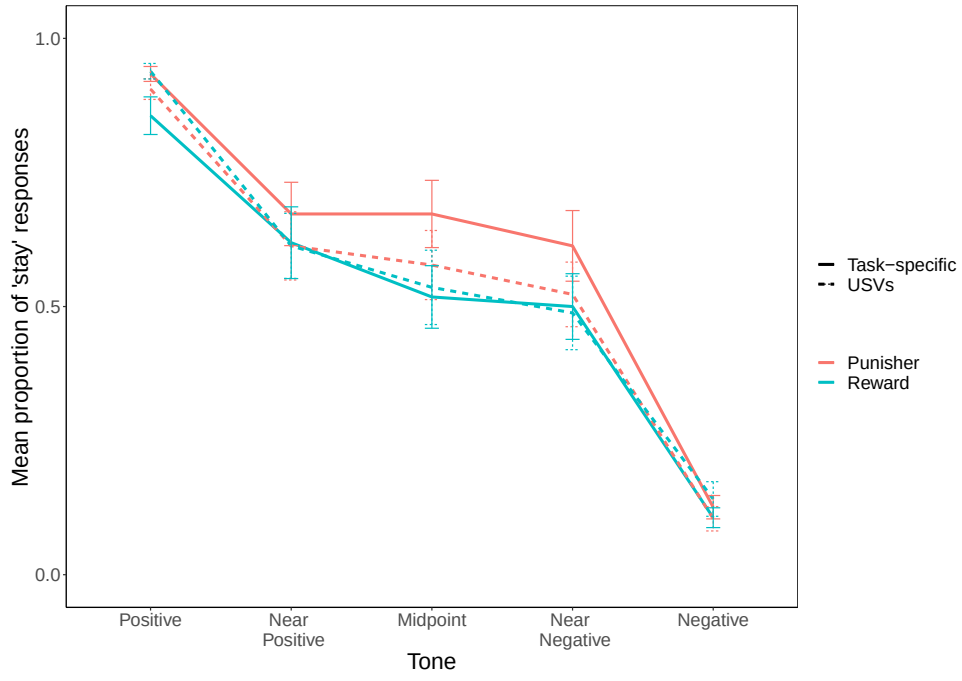


FIGURE 5.8: Mean proportion of risk-seeking responses split by tone presented following experience of rewards or punishers that were either task specific or USVs. Error bars represent one standard error

5.3.3 Model-dependent analysis

Each model was fitted using a three-level hierarchical Bayesian random effects analysis with a single top level empirical prior distribution of the model parameters across all subjects and within each subject. Therefore, the model fitting procedure was blind to the existence of the different conditions. We subsequently assessed the association between the posterior parameters and the conditions.

Judgement bias data

The Bayesian-decision theoretic model that was fitted to the judgement bias data included four parameters: bias (δ), differential sensitivity to the outcome ($C_{p/r}$), a slope parameter (σ) and a lapse rate (λ). According to the AIC values, the most parsimonious model included only two of these parameters ($C_{p/r}$ and σ); inclusion of δ λ and resulted in a less parsimonious model (Table 5.2).

TABLE 5.2: ΔAIC scores for computational models of judgement bias choice data: comparing the AIC value from each model to the AIC value of the best model

Model Parameters	**Δ AIC**
$C_{p/r}$, σ	0
δ,$C_{p/r}$, σ	20.93
$C_{p/r}$, P_{lapse}, σ	20.98
δ, $C_{p/r}$, P_{lapse}, σ	46.228
δ,$C_{p/r}$, P_{lapse}	84.547
$C_{p/r}$, P_{lapse}	104.463
δ, P_{lapse}	108.931
δ, σ	165.222
P_{lapse}	356.375
σ	356.38
P_{lapse}, σ	376.198
δ	25049.2
$C_{p/r}$	25054.2
δ,$C_{p/r}$	25068.36

Across all conditions and subjects, the log-transformed parameter estimates of $C_{p/r}$ (PT, mean $\pm$ SE: -0.143$\pm$0.047, p=0.003) were significantly less than zero, suggesting that general risk-seeking behaviour observed in the rats during the judgement bias task is attributed to greater sensitivity to rewards relative to punishers. While the estimates of $C_{p/r}$ did not vary significantly depending on the manipulation valence (LRT=2.054, p=0.253), prevalence (LRT=0.504, p=0.597), or modality (LRT=0.012, p=0.912), or the interaction between prevalence and valence (LRT=2.712, p=0.249), there was a significant interaction between modality and valence (Fig. 5.9: LRT=8.456, p=0.018). This significant interaction was driven

by the difference in parameter estimates between the pre-test sucrose and pre-test air-puff treatment groups, with significantly lower estimates of $C_{p/r}$ for the air-puff group (air-puff, mean±SE=-0.280±0.085; sucrose: mean±SE=0.001±0.104;LRT=-3.135, p=0.007) while the estimates of $C_{p/r}$ did not differ significantly between the 22kHz USVs and 50kHz USVs (LRT=0.981, p=0.695), 22kHz USVs and air-puff (LRT=-1.010, p=0.676), or 50kHz USVs and sucrose (LRT=-2.154, p=0.102) treatment groups. The estimates of σ did not depend on valence (LRT=1.208, p=0.453), prevalence (LRT=2.104, p=0.453), or modality (LRT=0.564,p=0.566), or the interactions between prevalence and valence (LRT=0.203, p=0.652) or modality and valence (LRT=1.457, p=0.453)

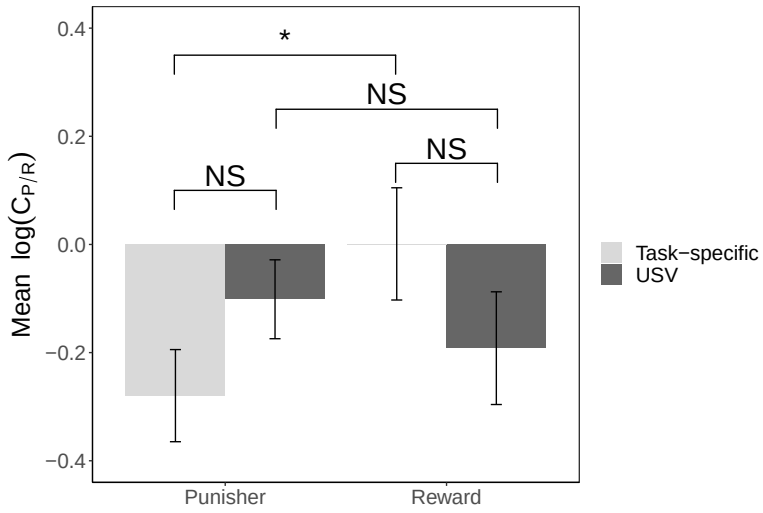


FIGURE 5.9: The mean parameter estimates of $C_{p/r}$ following experience of rewards or punishers that were either task specific or USVs Error bars represent one standard error

Self-determined inter-trial interval data

The linear regression model included a constant (k), which accounted for individual differences in the absolute speed of trial initiation, and regression weights for the previous outcome ($\beta_{R_{t-1}}$), weighted prediction error from the previous trial ($\beta_{\mathrm{PE}_{t-1}}$), squared weighted prediction error from the previous trial ($\beta_{\mathrm{PE}^2_{t-1}}$), and number of trials completed (β_t), as well as parameters σ and $C_{p/r}$ which described the expected outcome on each trial and hence, along with the actual outcome, determined the prediction error and scaled the previous outcome. The fitted learning rate for the average earning rate (mean±SE=0.794±0.014) indicated that the average earning rate largely reflected the most recent outcome. As the previous outcome provided a more parsimonious fit of the self-determined inter-trial intervals than the average earning rate, the average earning rate was not included in the model (ΔAIC=17.994; comparing all models including the average earning rate but not the previous outcome, to the same models instead including the previous outcome but not the average earning rate). As a result of the high degree of correlation between the weighted prediction

error, squared prediction error, and previous outcome the parameter estimates of $\beta_{\mathrm{PE}^2_{t-1}}$, $\beta_{\mathrm{PE}_{t-1}}$, and $\beta_{R_{t-1}}$ were highly correlated when fitted simultaneously. Models including both these regression weights were hence excluded from model comparison. The most parsimonious model according to the AIC value included k, $\beta_{R_{t-1}}$, β_t, and $C_{p/r}$ (Table 5.3).

TABLE 5.3: ΔAIC scores for computational models of self-determined inter-trial interval data: comparing the AIC value from each model to the AIC value of the best model

Model parameters	**ΔAIC**
k, $\beta_{R_{t-1}}$, β_t, $C_{p/r}$	0.000
k, $\beta_{\mathrm{PE}_{t-1}}$, β_t, $C_{p/r}$	3.002
k, $\beta_{R_{t-1}}$, $C_{p/r}$	119.539
k, $\beta_{\mathrm{PE}_{t-1}}$, $C_{p/r}$	122.601
k, $\beta_{\mathrm{PE}^2_{t-1}}$, β_t	135.366
k, $\beta_{R_{t-1}}$	144.216
k, $\beta_{\mathrm{PE}_{t-1}}$	146.045
k, $\beta_{\mathrm{PE}^2_{t-1}}$, $C_{p/r}$	177.077
k, $\beta_{\mathrm{PE}_{t-1}}$, γ_{wpe}	194.291
k, $\beta_{\mathrm{PE}^2_{t-1}}$	211.824
k, β_t	225.492
k, $\beta_{\mathrm{PE}^2_{t-1}}$, γ_{wpe}	255.943
k	336.645
δ,$C_{p/r}$	25068.36

Overall, the parameter estimates for $\beta_{R_{t-1}}$ (Fig 5.10: PT, 0.126±0.039, p=0.001), and $C_{p/r}$ (Fig 5.10: PT, 0.667±0.055, p<0.001) were found to differ significantly from zero. Rats were faster to initiate trials when the outcome of the most recent trial was more aversive and were more sensitive to losses than rewards. The estimates of β_t did not differ significantly from zero (Fig 5.10: p=0.616). With the exception of the estimates of $C_{p/r}$ which tended to be higher in the low prevalence compared with high prevalence condition (LRT=5.490, p=0.096), the parameter estimates did not significantly depend on manipulation valence, modality, prevalence, or the interactions between manipulation valence and modality, and manipulation valence and prevalence (Table 5.4).

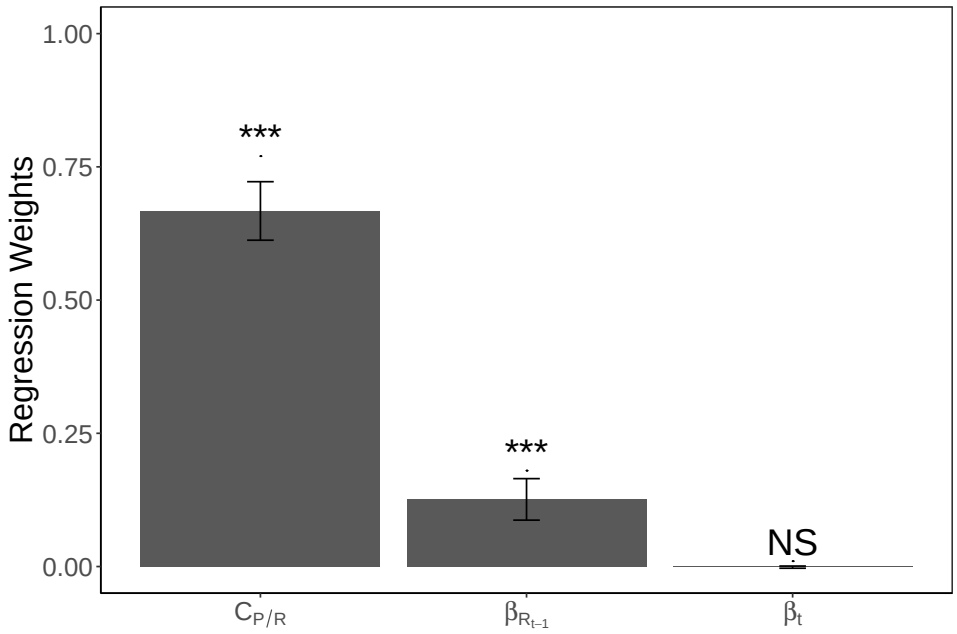


FIGURE 5.10: The mean estimate of regression weights $\beta_{R_{t-1}}$, and β_t and estimate of $C_{p/r}$. Error bars represent one standard error

TABLE 5.4: The LRT values and p-values from the statistical analysis of the parameter estimates from the model of self-determined inter-trial intervals

Parameter	**Predictor variable**	**LRT**	**p-value**
k	Valence:prevalence	0.749	0.536
	Valence:modality	2.102	0.368
	Valence	0.458	0.536
	Prevalence	0.384	0.536
	Modality	3.359	0.334
$\beta_{R_{t-1}}$	Valence:prevalence	0.011	0.979
	Valence:modality	0.050	0.979
	Valence	0.001	0.979
	Prevalence	0.004	0.979
	Modality	0.088	0.979
$C_{p/r}$	Valence:prevalence	0.312	0.988
	Valence:modality	0.061	0.988
	Valence	0.104	0.988
	Prevalence	5.490	0.096
	Modality	0.000	0.988
β_t	Valence:prevalence	1.482	0.279
	Valence:modality	2.094	0.247
	Valence	4.272	0.138
	Prevalence	0.160	0.690
	Modality	3.675	0.138

5.4 Discussion

Risky decision-making in the judgement bias task is considered to reflect a relatively positively valenced affective state (Bateson, 2016; Mendl et al., 2009). However, in our study, rats who had consumed sucrose prior to testing were more risk-averse than those who had received air-puffs, and rats worked less hard for sucrose in the lever pressing task when they had received pre-test rewards rather than pre-test punishers, irrespective of modality. Moreover, both these effects depended on the prevalence of the prior rewards or punishers, being significant only in the case of highly prevalent pre-test affective stimuli. Thus, although as hypothesised, judgement bias was modulated by the extent to which an individual's experience was rewarding or punishing, our results do not support the hypothesis that reward experience necessarily induces a more positively valenced affective state and consequently more risky decision-making relative to punisher experience. Instead, our results may reflect either that rewards are more valuable in negative affective states, or that the contrast from the negative affect induction to the test induced a relatively positive affective state.

Reward and punisher experience may indeed have induced a relatively more positive and negative affective state respectively, but expressed this in a way we had not originally hypothesised. An individual experiencing a negatively valenced affective state could be more risk-seeking if they valued rewards more highly or punishers less negatively. Rats that received pre-test punishers were more vigorous in their lever pressing in the progressive ratio-lever pressing task compared with those that received pre-test rewards. This suggests that pre-test punisher experience increased the subjective value of the sucrose pellet relative to pre-test reward experience. This is further supported by the results of the model-dependent analysis of data from the judgement bias task which revealed that rats that had experienced prior sucrose weighted present within-test punishers more heavily relative to within-test rewards than those that had experienced prior air-puffs.

Although it is arguably atypical, with anhedonic depression being more commonly observed, greater reward valuation has been observed in individuals experiencing more negative affective states. For example, depression in humans has been associated with increased valuation of a number of rewards including food (Simmons et al., 2016), cigarettes (Spring et al., 2003), and alcohol (Murphy et al., 2013). Similarly, animal studies have demonstrated that short-term stress can result in increased reward valuation (Spruijt et al., 2001; Van der Harst et al., 2003). It has been suggested this might function to regulate mood by encouraging individuals in poorer mood states to seek mood-enhancing rewards (Morris and Reilly, 1987; Sanchez et al., 2014). Additionally, Trimmer et al. (2017) suggested that food should be valued more highly in environments in which threats are more frequent as maintaining the high energy reserves needed to evade predation becomes more important (Trimmer et al., 2017).

An alternate explanation is that our results reflect a contrast effect. A few other studies have found an 'optimistic' judgement bias in individuals in affective states that

were putatively negative (e.g. Briefer and McElligott, 2013; Doyle et al., 2010; Sanger et al., 2011). A commonality between our study and some of these other studies is the release of individuals from an aversive situation prior to testing. In our study the putative affect manipulation occurred in a different location to the judgement bias task, and thus the context changed between the prior experience of rewards and punishers and the present test. As demonstrated by contextual fear conditioning, the context of an environment becomes associated with emotionally-salient experiences (Bouton and King, 1983; Hall and Honey, 1989). Therefore, it is possible that rats experienced a relief-like state following removal from the environment in which they had been exposed to punishers, and a disappointment-like state following removal from the environment in which they had experienced rewards.

This would be consistent with the successive negative and positive contrast effect observed in rats and other species; the disappointment-like or elation-like response observed when an individual experiences a downshift or upshift in the availability of a food reward (Burman et al., 2008b; Flaherty, 1982; Neville et al., 2017). Similarly, research with human subjects has demonstrated that self-reported affect is modulated by the difference between actual and expected earnings (reward prediction error) and not the absolute amount earned (Rutledge et al., 2014). The finding that reward sensitivity was enhanced following pre-test air-puff compared to pre-test sucrose experience would also be consistent with the contrast induced affective changes proposed here. Numerous studies have identified an association between negative affect and a decreased reward valuation (Huys et al., 2013; Treadway et al., 2012). This explanation would imply that use of environmental priors is context-dependent. In particular, an individual may recognize that prior experience might not be relevant in new contexts, and instead rely on the comparison of conditions between contexts to inform behaviour.

This study also revealed that decision-making depended on the modality of prior experience. While there was a difference in judgement bias between rats exposed to task specific rewards and punishers (i.e. pellets and air-puffs), contrary to the findings of Saito et al. (2016) in which a difference in judgement bias was observed between rats that were presented with 22kHz and 50kHz USVs, there was no difference between rats exposed to the rewarding and punishing USVs. The simplest explanation for this result is that the USVs were not as rewarding or punishing as the sucrose and air-puff, and consequently exerted a weaker effect on judgement bias. Additionally, unlike the study by Saito et al. (2016), the rats in this study were not housed in social isolation which may perhaps, as a result of increased exposure to USVs and social interaction, have reduced the salience of the USVs. However, this would be incongruent with the findings from the progressive lever ratio task in which the effect of manipulation valence was not found to depend on manipulation modality. An alternative or additional explanation could be that the prediction of outcomes in the judgement bias task was influenced by specific prior experience. The possibility that prior reward or punisher experience may have both general and specific effects on decision-making is

supported by human neural imaging studies which have demonstrated that there is only partial overlap of prediction error signals during learning with different reward and punisher modalities (Metereau and Dreher, 2012; Valentin and O'Doherty, 2009). The observed interaction between manipulation valence and modality thus raises the question of whether judgement bias is sensitive to the modality of the affect manipulation. Although numerous studies have found that affect manipulations unrelated to the task influence judgement bias, further investigation is needed to determine to what extent the congruity of prior experience to within-task decision-making modulates judgement bias.

Reward and punisher experience within the test session also altered behaviour. Rats were more likely to make the risky response when the previous outcome was more aversive. This could indicate that rats are able to more easily recall the most recent tone and compare whether its frequency is higher or lower than the current tone to deduce the likely outcome. It might also reflect a process similar to the tilt-shift effect (Gibson and Radner, 1937) whereby recently perceived stimuli alter the perception of subsequent stimuli.

The previous outcome also influenced the self-determined inter-trial intervals; rats were faster to initiate a trial when the previous outcome was more aversive. This finding could simply be attributed to the time spent consuming the sucrose pellet hence increased inter-trial intervals following rewarded trials. There was no evidence to suggest that the pre-test reward and punisher experience had an overall influence on self-determined inter-trial interval; the constant parameter in the model of self-determined inter-trial intervals which characterised the overall speed of trial initiation for each test session did not vary depending on pre-test experience suggesting that there was no detectable effect of the reward and punisher pre-treatments on latency to initiate trials.

In contrast to the model of judgement bias, the estimate of relative sensitivity to rewards and punishers in the model of self-determined inter-trial intervals suggested that rats were overall more sensitive to punishers than rewards. More specifically, given the structure of the model this finding indicates that experiencing an air-puff on the previous trial had a greater impact than sucrose on self-determined inter-trial intervals than the delivery of sucrose. This tended to be more pronounced in the loss condition. Although the reasons for this are unclear, it could speculatively reflect differences in the influence of potential and actual rewards on behaviour. The parameter in the model of judgement bias task reflected sensitivity to potential rewards and punishers, whereas in the model of inter-trial interval the parameter reflected sensitivity to recently experienced rewards and punishers.

One could argue that pre-test sucrose consumption could lead to satiation and accordingly a reduction of sucrose valuation which would result in risk-averse decision-making independent of changes in affect. However, although satiation could provide a partial explanation for our results, rats will consume far more sucrose pellets within a similar time period given free availability (Jones et al., 2018). Moreover, if satiation

did underlie our results it might be expected that rats would be more risk-averse as the session progressed and more sucrose was consumed, but our analysis found no time-dependent changes in judgement bias. Satiation would also not explain why the observed difference was between rats that had experienced the high and low rewards and punishers and also between rats that had experienced sucrose compared with air-puffs, as opposed to following the delivery of sucrose at a high frequency compared with all other pre-test reward and punisher conditions.

5.5 Conclusions

In summary, we found that (1) individuals were not more risk-seeking in the judgement bias task following reward compared with punisher experience but instead were more risk-averse; (2) this effect was dependent on reward and punisher prevalence with a greater effect found between rats that experienced the high compared to low prevalence rewards and punishers; (3) it was also dependent on the extent to which the modality of the rewards and punishers experienced matched that of the potential outcome of the decision, with a greater effect observed when the pre-test reward and punisher modality matched the within-test reward and punisher modality; (4) the observed judgement bias was found to result from a greater weighting of punishers relative to rewards of rats with prior experience of rewards compared with punishers as demonstrated by the results of the progressive ratio lever pressing task and model-dependent analysis; (5) the pre-test reward and punisher experience was not found to alter self-determined inter-trial intervals; however (6) vigour within the test session was modulated by the most recent outcome. The affective state induced by the manipulation is therefore unclear. The rats in the sucrose condition could have experienced a relatively positive affective state compared with the rats in the air-puff condition, with the observed shift in judgement bias reflecting an uncommon but nonetheless observed affect-related change in the subjective valuation of rewards or punishers. However, it is also possible that the contrast between the manipulation and test environment induced a relatively negative affective state in the rats that had experienced pre-test sucrose compared to pre-test air-puffs. This would suggest that affect not only depends on prior experience but also the context of such experience. This study therefore highlights that caution should be taken when interpreting risk-aversion as an indicator of relatively negative affect and relatively poor welfare in the judgement bias task. Although we can conclude that prior experience is a key determinant of vigour and decision-making, in future studies it will be important to investigate which aspects of the environment might modulate the influence of prior experience on behaviour.

Chapter 6

How does fluctuating reward magnitude influence rat judgement bias?

Chapter summary: It is common for affect manipulations to be conducted prior to behavioural tasks which attempt to measure the resulting affective state. However, interpretation of results using this methodology may be complicated by several confounding factors such as contrast effects or the affective impact of the behavioural task itself. To permit a better understanding of the relationship between reward and punisher experience, considered to govern affective state, and decision-making in rats, we conducted a judgement bias task in which we manipulated reward experience within testing by varying the volume of the potential apple juice reward on each trial. However, we encountered challenges in training rats on our novel variant of the judgment bias task and only four out of twenty rats completed the test session. Yet, even with this small sample size, we found that reward and punisher experience did influence behaviour. Rats were more likely to execute the risky action and were faster to initiate trials when they had learnt that the magnitude of apple juice received would be greater. Additionally, rats (although only one individual significantly) were more risk-seeking when they had been more successful on the task, which according the computational analysis specifically influenced decision-making via an increased reward sensitivity. Thus, while this study suggests that reward and punisher experience determines judgement bias in rats, further studies with a larger sample size will be required before more solid conclusions can be drawn.

6.1 Introduction

Investigating the relationship between affect and cognition typically involves conducting an affect manipulation prior to testing on a behavioural task, such as the judgement bias task. However, the task itself, or indeed the contrast between the affect manipulation and the task, might alter the subject's affective state and complicate interpretation of results. This could potentially explain some of the null or negative results from judgement bias studies (e.g. Briefer and McElligott, 2013; Burman et al., 2011; Doyle et al., 2010), including the finding that rats are more risk-seeking following experience of punishers compared to rewards (Chapter 5). Hence, there is a need for a better understanding of how within-test experience might influence decision-making, and also for methods of inducing negative or positive affective states during the test session.

Affect is proposed to reflect an individual's experience of rewards and punishers (Eldar et al., 2016; Mendl et al., 2010; Nettle, 2008). This is supported by the findings of Chapters 3 and 4 which demonstrated that the average earning rate influenced subjectively reported affective valence in humans, although other studies have suggested that the difference between an actual outcome and the expected outcome determines subjective affective valence (Rutledge et al., 2014). Additionally, using computational modelling as described in Chapters 3 and 4, we found that decision-making depended on how predictable recent outcomes had been, and that the extent to which predictability modulated decision-making was associated with affective valence. Examining the influence of within-test reward and punisher experience might therefore be a useful means of elucidating the relationship between reward and punisher experience and cognition. Moreover, parameters extracted from judgement bias data using a task using computational modelling could provide a useful and more reliable measure of affective valence.

In this study, we adapted the automated judgement bias task for rats described by Jones et al. (2018) to allow within-test variation of reward magnitude. Specifically, following the human judgement bias task described in Chapter 4, we varied the reward (also apple juice in this study) magnitude according to a noisy sine wave. The choice data from this task was analysed according to our novel model of judgement bias (Chapters 3 and 4) to assess the influence of reward and punisher experience on the putative processes underlying decision-making, namely reward and punisher sensitivity and prior beliefs about the reward and punisher. We hypothesised that rats would be more likely to make the risky response when the potential juice reward is higher, and that above this effect, reward and punisher experience would influence decision-making.

6.2 Methods

6.2.1 Subjects

Subjects were 20 male Lister Hooded rats (Charles River, Margate, UK), aged c.9 weeks of age when training commenced. Rats were housed in pairs in cages measuring 560x340x190mm under a 12-hour reversed light-dark cycle (lights on 1900-0700). Cages contained sawdust substrate, shredded paper bedding as well as a number of enrichment items including a cardboard nest box, a plastic tube attached to the cage lid, a cardboard tube, and an aspen block. Food (LabDiet) and water were available *ad libitum* and rats were not food or water restricted prior to training or testing. All rats were checked regularly for any health issues, and the work was conducted under University of Bristol Investigation Number UB/16/004 following approval by the Bristol University Animal Welfare and Ethics Review Body. All rats were rehomed as pets on completion of the study.

6.2.2 Apparatus (see Fig. 6.1)

Training and testing were conducted in a shuttle box (508×254×305mm) placed in a sound isolation chamber. The box was divided in half by a metal panel and only one half (254×254×305mm) was used throughout the entire study. A custom-made trough (Faculty of Engineering, University of Bristol) was located centrally on the end wall. An opening (32×40×35mm) which was 35mm above the floor allowed access to the trough from the shuttle box. Syringe pumps controlled the delivery of apple juice (Morrisons, West Yorkshire, UK; Apple Juice from Concentrate) to these troughs. A stainless steel tube was inserted through a hole in the side wall of the trough and was connected to a pump-controlled syringe via PVC tubing. All hardware was manufactured by Coulbourn Instruments (Allentown, PA, USA), and operated by their Graphic State (v4) software.

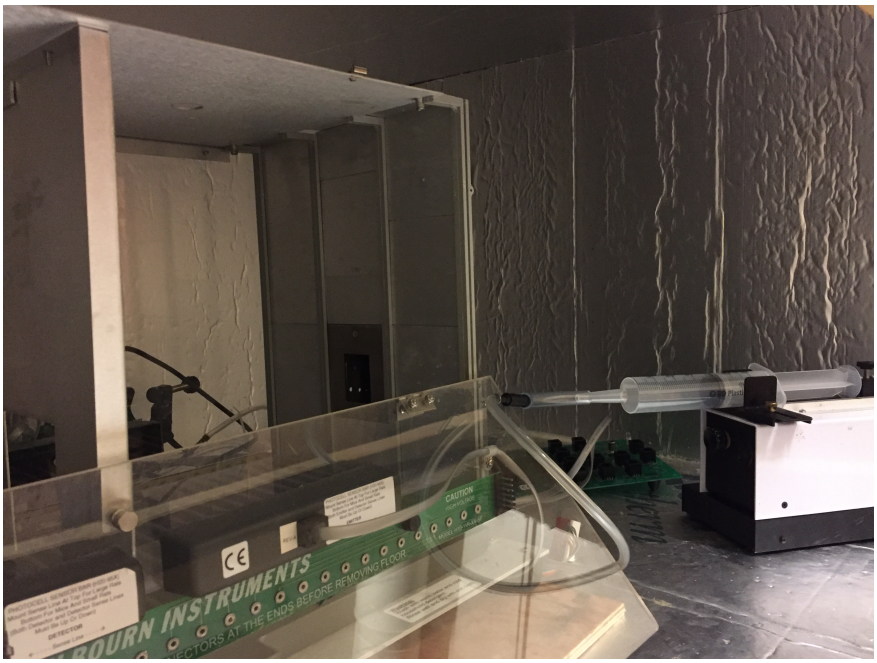


FIGURE 6.1: The shuttle box within a sound isolation chamber used for the judgement bias task; there is a syringe pump connected to the trough on the end wall via PVC tubing.

6.2.3 Training

Following Jones et al. (2018), rats were trained to 'stay' in a trough following presentation of a tone (either 2Khz or 8khz) to obtain a food reward, and to leave the trough ('go') following presentation of the alternate tone (either 2Khz or 8khz) to avoid a punisher. Hence, there is a safe response ('go') and a risky response ('stay'). In human versions of this task, subjects are informed of the potential reward and punisher magnitude prior to testing. To parallel this in the rat task, the amplitude of the tone varied according to the magnitude of the potential reward; higher amplitudes signalled a greater reward magnitude. Use of a liquid reward allowed fine-grained variation in reward magnitude (i.e. volume delivered) and we opted to use apple juice as the reward as we considered that sucrose solution would be too viscous and would create difficulties in cleaning the equipment. We verified that rats found apple juice rewarding and assessed the quantities they would consume in a pilot study. Due to equipment constraints, in particular that we could not vary the strength (psi) of air-puff on a trial-by-trial basis, the magnitude of the air-puff could not be varied within-testing but instead was fixed at 40psi. However, the amplitude of the negative tone was varied so that amplitude did not provide information about the potential outcome.

Rats were first habituated to the training shuttle boxes. This involved two 15 minutes sessions on consecutive days in which rats were placed in the shuttle box with no trough, on the first day each rat was accompanied by their cagemate and on the second day they were placed in the shuttle box on their own. Rats then completed one training session of 30-minute duration every weekday until they met the criterion

to progress to testing. Rats first underwent positive training in which they learnt to associate the positive tone with apple juice, the volume of which depended on the amplitude of tone (0.02 ml for the low amplitude tones and 0.1ml for the high amplitude tones), and to stay in the trough for increasing durations to receive this apple juice. In this phase, the high amplitude positive tone was presented on 33% of trough visits, the positive low amplitude tone was presented 33% of trough visits, and on the remaining 33% of visits no tone was presented (null tone). The order of these trials was randomised.

Rats were first trained to keep their snout in the trough for 20ms (Stage=p1) to receive the apple juice. If the rat visited the trough at least 40 times, stayed in the trough for the required time on 62.5% of trials in which the positive tone was presented, and completed at least two sessions, the required stay duration between trough entry and juice delivery was increased to 0.75s (Stage=p2). If the same criteria were met in this stage, the duration was increased to 1.5s (Stage=p3). An individual progressed to final stage of positive training according to the same criteria. In comparison to previous studies using this methodology, rats were required to keep their snout in the trough for 2s to receive apple juice during the final stage of positive training (Stage=p4). The purpose of this stage was to ensure that rats would wait for 2s to obtain the low volumes of juice. Rats progressed to discrimination training following the same criterion as before; 62.5% accuracy for each tone and at least 40 trials completed in a session, although were only required to complete one session of this stage.

Discrimination training involved training rats to 'stay' in response to the positive tones to obtain apple juice and to 'go' in response to the negative tones to avoid an air-puff. The positive and negative tone was counterbalanced across individuals. In the first stage of discrimination training (Stage=d1), the positive tone was presented on 50% of trials (25% high and low), negative tones on 25% of trials (12.5% high and low), and null tones on 25% of trials. The order of tone presentation was randomised. To progress to the next stage of discrimination training (Stage=d2), an individual was required to visit the trough at least 60 times within a session, make the correct response on 62.5% of visits for both the negative and positive tones, and have completed at least two sessions at this stage of training. In the second and final stage of discrimination training, the positive and negative tone were presented on 46% of trials each (23% high and low amplitude) and null tone on 8% of trials. If rats met the same criteria described above, they progressed to judgement bias testing. If any individual performed poorly (<50% accuracy) for four consecutive days, they returned to the previous stage of training.

However, few rats learnt to discriminate between the positive and negative tone, which was particularly apparent for the low amplitude tones. Hence, from the fourth week of training onwards, rats were only trained on the high amplitude tones. Despite this, we continued to have difficulties training rats and so reduced the strength of the air-puff to 20psi for poorly performing rats.

6.2.4 Testing

During the test session, in addition to the reference tones used in testing, rats were also presented with four different probe tones (2.639kHz, 3.482kHz, 4.595kHz, and 6.063kHz). The order of the tone presentation was randomised but identical for all rats. Tones which were perceptually closest to the positive reference tone were rewarded if the 'stay' response was made and tones which were perceptually closest to the negative tone were punished if the 'stay' response was made, nothing was received following a 'go' response. The volume of apple juice delivered for correct stay responses was pre-determined according to a noisy sine wave with mean=0.118ml and standard deviation=0.027ml, between a minimum volume of 0.067ml and maximum volume of 0.150ml (Fig. 6.2). There were 90 trials and each tone was presented 15 times each. If a rat did not complete sufficient trials within a test session (<50% trials completed), they returned to training for at least two consecutive days before repeating the test session.

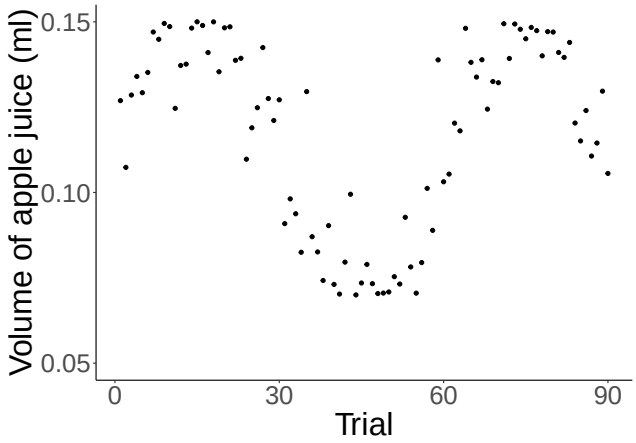


FIGURE 6.2: Offered volume of apple juice on each trial of the judgement bias task

6.2.5 Model-dependent analysis

The model-dependent analysis assessed the key hypothesis of this study; to investigate the cognitive processes underlying judgement bias, and how these relate to reward and punisher experience. In particular, we modelled judgement bias choice data as a partially observable Markov decision process (POMDP), as described in Chapters 3 and 4. Briefly, the model considers each rat to occupy a belief state that depends on the objective time and observations of the stimulus up to that time. Each belief state has a value that is defined by the likely transitions from that state given the current state, the potential reward and punisher, and state-dependent probability of the reward and punisher. The state value determines the probability that an individual will make either response upon reaching the state according to a softmax function. We can then compute the probability that a rat makes the 'go' response by considering both the probability that they transition to any state given the presented stimulus

and the probability of making the 'go' response in those states. In contrast to human participants who know the potential value of the reward and punisher, rats must learn the value of the potential reward according to experience. We consider that rats update their expected value of the reward (learnt reward; an average reward rate) following each delivered reward according to the Rescorla-Wagner learning rule (Rescorla et al., 1972) with learning rate (α). The prediction error ($w\mathrm{PE}_{n-1}$) and average earning rate ($\bar{R}_{n-1}$) also depend on previous outcomes as previously described in Chapter 3; the prediction error is the difference between the outcome and the model-determined stimulus-dependent predicted outcome (which fit the data better than when modelled using the prediction formed prior to stimulus onset, ΔAIC=56.726). The average earning rate ($\bar{R}_{n-1}$) is updated following each outcome according to the Rescorla-Wagner learning rule, and in contrast to the learnt reward also includes both punishing and null outcomes. In the model, biases in decision-making arise through an altered reward sensitivity, loss sensitivity, or the prior belief that the trial will be rewarded. The model assumes that $\bar{R}_{n-1}$, $w\mathrm{PE}_{n-1}$, $w\mathrm{PE}^2_{n-1}$, and O_{n-1} influence C_R, C_P, and ω additively, with β-values which determine the extent of their influence as follows:

$$\omega = e^{\left(\beta_0^{\omega}+\beta_{\bar{R}}^{\omega}\bar{R}_{n-1}+\beta_{w\mathrm{PE}}^{\omega}w\mathrm{PE}_{n-1}+\beta_{w\mathrm{PE}^2}^{\omega}w\mathrm{PE}^2_{n-1}+\beta_{O}^{\omega}O_{n-1}\right)} \tag{6.1}$$

$$C_R = e^{\left(\beta_0^{C_R}+\beta_{\bar{R}}^{C_R}\bar{R}_{n-1}+\beta_{w\mathrm{PE}}^{C_R}w\mathrm{PE}_{n-1}+\beta_{w\mathrm{PE}^2}^{C_R}w\mathrm{PE}^2_{n-1}+\beta_{O}^{C_R}O_{n-1}\right)} \tag{6.2}$$

$$C_L = e^{\left(\beta_0^{C_L}+\beta_{\bar{R}}^{C_L}\bar{R}_{n-1}+\beta_{w\mathrm{PE}}^{C_L}w\mathrm{PE}_{n-1}+\beta_{w\mathrm{PE}_n^2}^{C_L}w\mathrm{PE}^2_{n-1}+\beta_{O}^{C_L}O_{n-1}\right)} \tag{6.3}$$

In addition to these parameters, the following parameter were also assessed for their contribution to the model fit: σ which accounts for the rats' ability to discriminate between the tones, a lapse rate (λ) which accounts for stimulus independent errors, a joint learning rate for the average earning rate and average reward rate (α), and a forgetting factor for the weighted prediction error (γ). We took a stepwise approach to model fitting, and compared the goodness of fit of the models according to their AIC values. Model-fitting was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bris.ac.uk/acrc/, as described in Chapter 3.

6.2.6 Statistical analysis

Data were analysed using generalised linear mixed models (GLMMs) using the R (R Core Team, 2017) packages lme4 (Bates et al., 2015) and nlme (Pinheiro et al., 2018). All GLMMs included a random effect of subject.

Data from training sessions were analysed to assess progression and the factors influencing progression. We modelled the number of trials completed in each test

session from the stage of training (i.e. p1, p2, p3, p4, d1, and d2). Additionally, accuracy in the positive phase of training (i.e. p1:p4) was modelled from training stage and tone amplitude, accuracy in the first stage of discrimination training (d1, prior to removal of the low amplitude tones) from tone frequency and tone amplitude, and accuracy across discrimination training (i.e. d1 and d2) from stage and tone frequency. To make post-hoc comparisons of factor levels that were significant, we used Tukey's Honest Significant Difference (HSD) test using the R package multcomp (Hothorn et al., 2008). The statistical analysis of training data allowed us to assess the rats performance during training, to allow a better understanding of how training could be improved in future.

We also analysed both judgement bias (the binary decision to 'stay' or 'go') and the self-determined inter-trial intervals (the time between the end of one trial and initiation of the next trial). In the GLMM of judgement bias, fixed effects included the learnt reward (derived from the model-dependent analysis), the number of trials completed, and the stimulus presented. The GLMM of self-determined inter-trial interval also included the number of trials completed and the learnt reward, and also included $\bar{R}_{n-1}$, O_{n-1}, $w\mathrm{PE}_{n-1}$, and $w\mathrm{PE}^2_{n-1}$ (derived from the model-dependent analysis). To avoid issues of multicollinearity, O_{n-1}, $w\mathrm{PE}_{n-1}$, and $w\mathrm{PE}^2_{n-1}$ were not included in the same model but instead separate models including each variable were compared according to their AIC value. The best fitting model included the squared prediction error (ΔAIC=5.158, comparing saturated model containing $w\mathrm{PE}^2_{n-1}$ and not $w\mathrm{PE}^2_{n-1}$ or O_{n-1} with the next best-fitting saturated model which contained O_{n-1} but not $w\mathrm{PE}_{n-1}$ or $w\mathrm{PE}^2_{n-1}$). The statistical analysis of judgement bias data allowed us to assess the extent to which rats attended to the fluctuating reward, the extent to which they discriminated between the tones, and the effect of number of trials completed on judgement bias. The statistical analysis of self-determined inter-trial intervals allowed us to assess whether there were time and reward volume dependent changes in motivation to obtain juice.

6.3 Results

6.3.1 Training

Training of the rats during the positive phase was comparable to that described in both Jones et al. (2018) and Chapter 5; rats completed the first stage of positive training (p1) in 2.350±0.150 sessions (mean± SE), the second stage (p2) in 2.250±0.315 sessions, the third stage (p3) in 2.000±0.000 sessions, and the final stage (p4) in 2.368±0.738 sessions. All 20 rats completed positive training. The number of training sessions required before rats (who did meet the progression criterion) reached criterion in discrimination training was also comparable; rats progressed from the first to second stage of discrimination in an average of 7.643±1.112 sessions and from the second phase of discrimination to testing in 7.500±1.648 sessions. However, the proportion of rats who failed to reach criterion was much greater that described in

both Jones et al. (2018) and Chapter 5, with only 15 rats (75%) completing the first stage of discrimination training (five did not progress due to both insufficient trials and inaccuracy) and only six (30%) progressing to testing (five did not progress due to insufficient trials only, three due to both insufficient trials and inaccuracy, and one due to inaccuracy). Although six rats progressed to testing, only four completed a sufficient number of trials for their data to be analysed.

Training stage was a significant predictor of accuracy in the positive training phase (Fig. 6.3: LRT=103.812, $p<0.001$). More specifically, accuracy decreased significantly across consecutive stages except between p3 and p4 (Table 6.1). Accuracy did not differ significantly between the high and low amplitude tones (LRT=0.841, p=0.359), and there was no significant interaction between training stage and tone amplitude (LRT=3.633, p=0.304).

TABLE 6.1: z-values and p-values from the post-hoc contrast analysis of the influence of training stage on accuracy during the positive training phase

Contrast	**z-value**	**p-value**
p2 vs. p1	-3.232	0.007
p3 vs. p1	-10.211	<0.001
p4 vs. p1	-7.074	<0.001
p3 vs. p2	-5.588	<0.001
p4 vs. p2	-3.703	0.001
p4 vs. p3	1.233	0.603

In the first stage of discrimination training (only considering data prior to the removal of the low amplitude tones) rats were both less accurate following presentation of the positive tones compared to the negative tones (Fig. 6.3: LRT=140.777, $p<0.001$) and also following presentation of the low amplitude tones compared to the high amplitude tones (Fig. 6.3: LRT=106.760, $p<0.001$). Furthermore, there was a significant interaction between tone amplitude and tone frequency (LRT=18.821, $p<0.001$). Post-hoc analysis revealed that rats had the greatest accuracy when the high amplitude negative tone was presented, the lowest accuracy when the low amplitude positive tone was presented, and there was no significant difference in accuracy following presentations of the low amplitude negative tone and high amplitude positive tone (Table 6.2).

Rats had a greater accuracy in the second stage of discrimination training compared with the first stage (Fig. 6.3: LRT=49.011, $p<0.001$), and also following presentation of negative tones compared to positive tones (Fig. 6.3: LRT=190.103, $p<0.001$). There was no significant interaction between stage and tone frequency (LRT=1.889, p=0.169).

TABLE 6.2: z-values and p-values from the post-hoc contrast analysis of the interaction between amplitude and valence on accuracy during the first stage of discrimination training

Contrast	**z-value**	**p-value**
Low -ve vs. high -ve	-4.874	<0.001
High +ve vs. high -ve	-8.593	<0.001
Low +ve vs. high -ve	-17.969	<0.001
High +ve vs. low -ve	-1.864	0.24
Low +ve vs. low -ve	-11.077	<0.001
Low +ve vs. high +ve	-11.183	<0.001

FIGURE 6.3: Mean proportion of accurate responses to presented stimuli split by tone and amplitude in each stage of training

The number of trials initiated varied significantly depending of the stage of training (Fig. 6.4: LRT=113.021, p<0.001). This was driven by an increase in the number of trials initiated in the final three stages of positive training, compared with the first stage of positive training and both stages of discrimination training (Table 6.3).

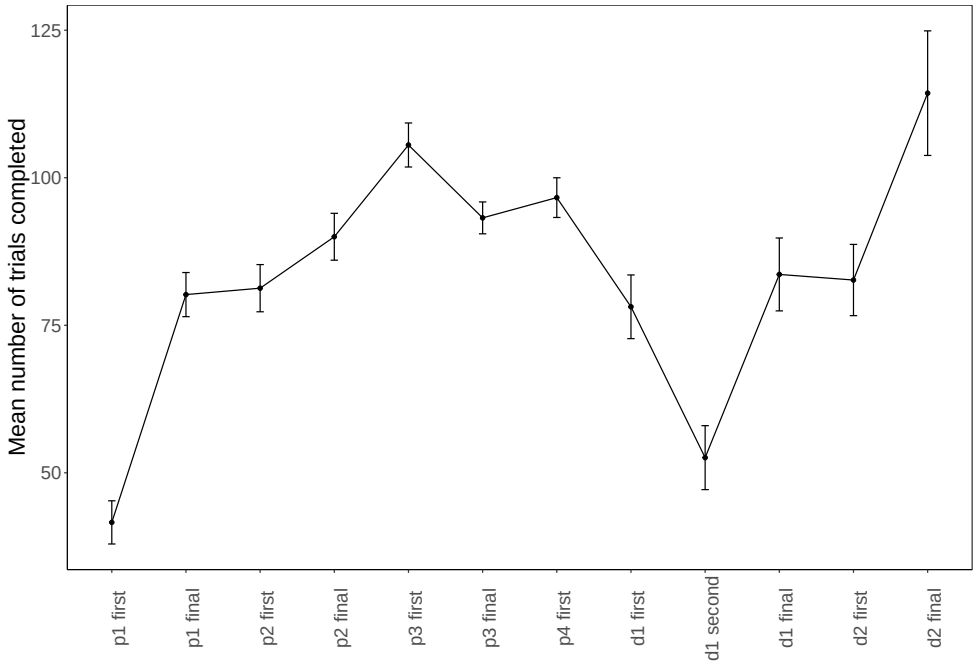


FIGURE 6.4: Mean number of accurate trials completed in each stage of training

TABLE 6.3: z-values and p-values from the post-hoc analysis of the influence of training stage on number of trials completed

Contrast	**z-value**	**p-value**
d2 - d1	-0.561	0.993
p1 - d1	0.543	0.994
p2 - d1	6.801	<0.001
p3 - d1	6.684	<0.001
p4 - d1	5.945	<0.001
p1 - d2	1.018	0.908
p2 - d2	7.021	<0.001
p3 - d2	6.862	<0.001
p4 - d2	6.209	<0.001
p2 - p1	6.32	<0.001
p3 - p1	6.079	<0.001
p4 - p1	5.476	<0.001
p3 - p2	-1.381	0.73
p4 - p2	-0.231	0.999
p4 - p3	1.016	0.909

6.3.2 Testing

The model-dependent analysis revealed that the best fit model included the following parameters: σ, $\beta_0^{C_R}$, $\beta_{\bar{R}}^{C_R}$ (Table 6.4, Fig. 6.5). A variable joint learning rate for

the average reward rate and average earning rate α was found to improve model fit (ΔAIC=55.011). The model derived probability of making a 'stay' response was found to be a strongly significant predictor of the observed response when analysed using a binomial GLMM with a random effect of subject (Fig. 6.6: LRT=186, p<0.001)

TABLE 6.4: ΔAIC values for computational models of judgement bias choice data: comparing the AIC value for each model with the AIC value of the best model

Model Parameters	**ΔAIC**
σ, $\beta_0^{C_R}$, $\beta_{\bar{R}}^{C_R}$	0.000
σ, $\beta_0^{C_R}$	1.553
σ, $\beta_0^{C_R}$, $\beta_{w\mathrm{PE}}^{C_R}$	9.486
σ, $\beta_0^{C_R}$, $\beta_{w\mathrm{PE}^2}^{C_R}$	20.239
σ, $\beta_0^{C_R}$, $\beta_{O_{n-1}}^{C_R}$	20.303
σ, β_0^{ω}, $\beta_{w\mathrm{PE}^2}^{\omega}$	171.623
σ, β_0^{ω}, $\beta_{w\mathrm{PE}}^{\omega}$	172.568
σ, β_0^{ω}, $\beta_{O_{n-1}}^{\omega}$	172.818
σ, β_0^{ω}	176.081
σ, β_0^{ω}, $\beta_{\bar{R}}^{\omega}$	180.054
σ, $\beta_0^{C_L}$, $\beta_{w\mathrm{PE}^2}^{C_L}$	343.586
σ, $\beta_0^{C_L}$,$\beta_{w\mathrm{PE}}^{C_L}$	344.139
σ, $\beta_0^{C_L}$	398.435
σ, $\beta_0^{C_L}$, $\beta_{\bar{R}}^{C_L}$	404.629
σ, $\beta_0^{C_L}$, $\beta_{O_{n-1}}^{C_L}$	415.925
σ	810.431
σ, λ	818.431
λ	877.250

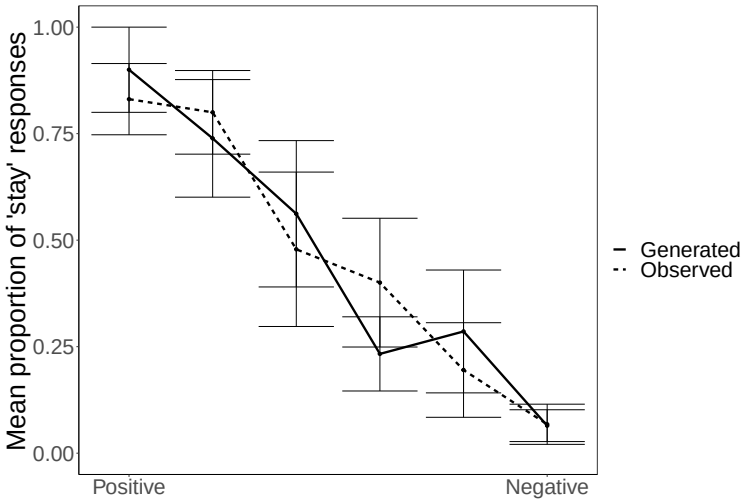


FIGURE 6.5: The mean proportion of 'stay' responses from the model-generated and observed judgement bias data. Errors bars represent one standard error

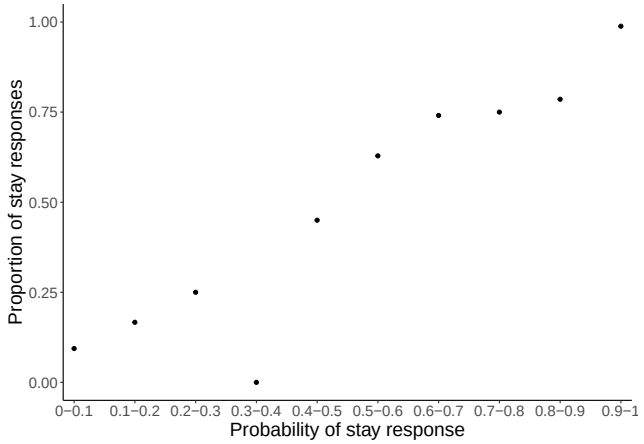


FIGURE 6.6: The proportion of 'stay' responses for intervals of model-derived probabilities of executing the 'stay' response

As the sample size (four rats) was insufficient to run statistical tests as in Chapters 3 and 4, we instead computed a 95% confidence interval for each parameter estimate as follows:

$$\text{Upper 95\% limit} = p + 1.96\sqrt{\phi^{-1}} \tag{6.4}$$

$$\text{Lower 95\% limit} = p - 1.96\sqrt{\phi^{-1}} \tag{6.5}$$

where ϕ^{-1} is the inverse hessian around parameter estimate (p) obtained from the model fitting procedure.

Although all estimates of $\beta_0^{C_R}$ were positive, indicative of a greater sensitivity to the apple juice relative to the true volume of juice, this was not reliable for any of the rats (Table 6.5). Similarly, although estimates of β^{C_R} were all greater than zero,

indicative of a greater reward sensitivity when the average earning rate was higher, this was reliable in only one of the four rats (Table 6.5). The mean (±SE) value of α was 0.130±0.100.

TABLE 6.5: Parameter estimates of $\beta_0^{C_R}$ and $\beta_{\bar{R}}^{C_R}$ for each rat with the upper and lower 95% confidence intervals from the computational model of judgement bias choice data

Parameter	**Rat**	**Estimate**	**Lower 95% CI**	**Upper 95% CI**
$\beta_0^{C_R}$	1	1.2843	-5.269	7.837
	2	0.9855	-2.279	4.250
	3	1.6659	-291.132	294.463
	4	0.4438	-0.041	0.929
$\beta_{\bar{R}}^{C_R}$	1	0.4184	-4.058	4.894
	2	0.0067	-2.310	2.323
	3	0.0181	-30.066	30.102
	4	4.1085	1.430	6.787

The stimulus was a significant predictor of judgement bias, indicating that rats did learn to discriminate between the tone frequencies (LRT=131.452, p<0.001). The learnt reward was also a significant predictor of judgement bias (Fig. 6.7: LRT=14.687, p<0.001); rats were more risk-seeking when the learnt reward was higher. The number of trials completed did not significantly influence judgement bias (LRT=1.223, p=0.269).

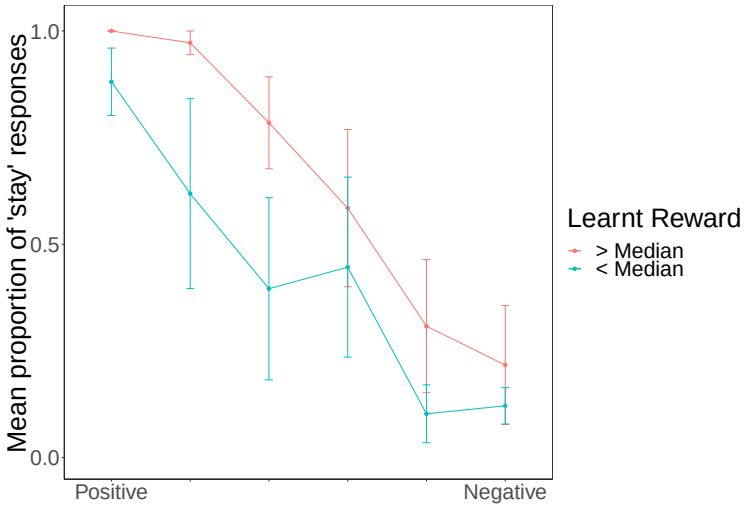


FIGURE 6.7: Mean proportion of 'stay' responses for each presented stimulus when the learnt reward was greater than and less than the median value. Error bars represent one standard error

Rats were faster to initiate trials when the learnt reward was higher (LRT=4.364, p=0.037) and when $w\text{PE}_{n-1}^2$ (squared weighted prediction error) was lower (LRT=7.182, p=0.007). $\bar{R}_{n-1}$ (the average earning rate; LRT=0.075, p=0.785) and

number of trials completed (LRT=1.708, p=0.191) did not significantly influence the self-determined inter-trial interval.

6.4 Discussion

The aim of this study was to develop a variant of the automated judgement bias task for rats in which reward availability varied across testing, allowing us to investigate how within-test fluctuations in reward magnitude, assumed to influence affective state, altered decision-making behaviour. To achieve this, we rewarded rats with apple juice and varied the volume delivered for correct 'stay' responses according to a noisy sine wave across judgement bias testing.

Unfortunately, few rats met the progression criteria for testing. Analysis of the training data indicated that problems arose when the negative tone was introduced, and that there were particular issues with the low amplitude and positive tones. Overall, rats were better at making the 'go' response for the negative tone than the 'stay' response for the positive tone. This suggests that following the introduction of the air-puff, the apple juice, in particular the very small volume delivered following the low amplitude tone, may have been insufficiently rewarding such that rats were unwilling to risk an air-puff to obtain it. This could perhaps be solved with food or water restriction, or using an alternate liquid that may be more rewarding without leading to satiation, for example sucrose solution could be a potential alternative reward (if potential issues of equipment cleaning were solved).

Rats were also less accurate at the low amplitude tone for both the positive and negative tones. This could indicate that it is more difficult for rats to determine the frequency of the tone at low amplitudes, given the lower signal to noise ratio, leading to uncertainty about whether a reward or punisher would be delivered and hence leaving them unable to learn the correct response. However, although a high proportion of rats were able to learn the discrimination following removal of the low amplitude tones, few rats met the criterion to progress to testing. This might reflect that negative tones were more frequently presented in the second stage of discrimination training, resulting in an increased unwillingness to initiate trials.

However, the rats that progressed to testing demonstrated that they were attending to within-test reward experience and modulating their behaviour accordingly. Rats were more likely to make the risky 'stay' response when they had learnt that juice availability was higher. Thus, consistent with results from the human version of the task and theoretical work (Bernoulli, 1954; Daston, 1980; Ore, 1960; Stearns, 2000; Von Neumann and Morgenstern, 1953), the expected value of the trial is an important determinant of the response made.

The model-dependent analysis suggested that judgement bias was modulated by the average earning rate. However, although the parameter estimates were associated with greater reward sensitivity and hence greater risk-seeking when the average earning rate was higher, this was only reliable for one of the four rats. Further research

with a larger sample size should be conducted to assess the extent to which a high average earning rate induces a relatively positive judgement bias in rats. Importantly, in humans (see Chapter 3 and 4), a higher average earning rate was associated with more positive affective valence. Although we cannot conclusively determine whether rats subjectively experience positive or negative affective states, we can state that the aspects of reward experience that modulate momentary subjective affective valence in humans also inform decision-making in rats.

In accordance with the results of the human task using juice rewards (Chapter 4), the extent of the judgement bias could be attributed to variation in reward sensitivity. Although the model-estimated parameters reflected that rats valued the juice more highly than expected given the true volume of the juice, which here is perhaps self-selecting (i.e. only the rats that strongly valued the juice were sufficiently motivated to learn the discrimination and complete trials, and so meet the progression criteria for testing), this was not reliable in any of the four rats.

The results of the analysis of inter-trial interval data was also comparable to the results of the human primary reward and punisher judgement bias study (Chapter 4). Rats (and humans) were slower to initiate trials when recent outcomes had been less predictable. This might reflect that individuals take time to consolidate new information or to seek additional cues that may aid the reduction of future unpredictability. The learnt reward, essentially an average reward rate which differs from the average earning rate in that it only tracks rewards and not null or punishing outcomes, also influenced the self-determined inter-trial interval. Rats were faster to initiate trials when they had learnt that the magnitude of the available reward was greater, perhaps allowing them to capitalise on highly rewarding periods of the test sessions (Beierholm et al., 2013; Niv et al., 2007).

6.5 Conclusions (see Fig. 6.8)

Although it is clear that our novel variant of the judgement bias task could be used to investigate the influence of reward experience on decision-making without the complexities of manipulating experience prior to testing, there are methodological issues that must be first addressed. Alternate liquid rewards should be sought which have a greater incentive salience than apple juice. Yet, despite the problems during training, the results from the rats who progressed to testing were promising. In particular, we demonstrated that reward and punisher experience does modulate judgement bias and vigour. It would therefore be worthwhile to attempt to remedy the methodological issues and repeat this study.

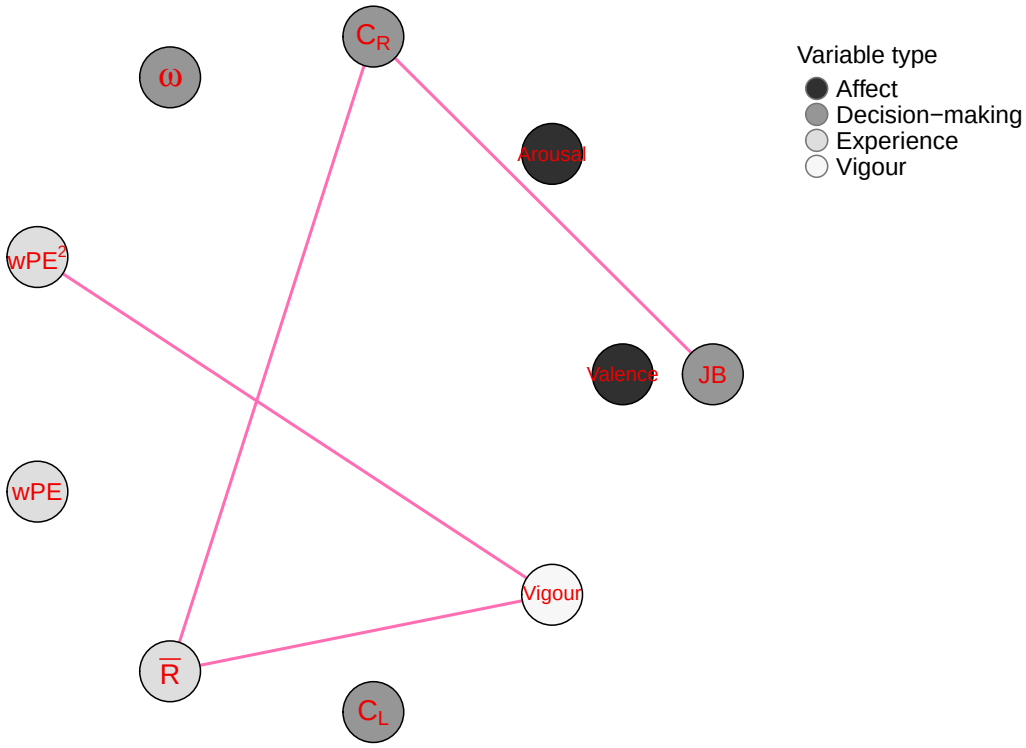


FIGURE 6.8: Diagrammatic summary of the results of Chapter 6: the nodes represent variables relating to affect, decision-making (where C_L denotes loss/punisher sensitivity, C_R denotes reward sensitivity, ω denotes the prior belief about outcomes, and JB denotes judgement bias), experience (where wPE denotes the weighted reward prediction error, $w\text{PE}^2$ denotes the squared weighted reward prediction error, and $\bar{R}$ denotes the average earning/reward rate), and vigour, the lines between nodes represent observed associations between those variables.

Chapter 7

Affective state and exploratory behaviour in rats

Chapter summary: Negative affective states in humans (particularly depression) have been shown to alter exploratory behaviour. Here, we assess how exploration might relate to affect in rats and whether a task which measures exploration could provide a novel measure of affect that requires less training than the judgement bias task. The task developed in this study required rats to choose between persevering with a depleting food source or forgoing a food reward to move to a new food source. Data were analysed using a computational model based on the marginal value theorem which considered the optimality of exploration relative to exploitation, and using a model-agnostic statistical analysis. We found that rats in a putative negative affective state, induced by the removal of enrichment from the homecage, stayed at the depleting food source for fewer consecutive trials (greater exploration) and completed fewer trials in a test session. The computational analysis revealed that following the removal of enrichment rats exhibited a reduced rate of learning and increased stochasticity in decision-making. However, equipment failure reduced the sample size of this study to six rats and so further studies will be required before conclusions can be drawn about the suitability and reliability of this task as a measure of affect in non-human animals.

7.1 Introduction

It is self-evident that successful foraging is vital to survival. However, food sources are typically not replenished at the same rate at which they are consumed. This leaves organisms with a predicament: to persist with a depleting resource or to expend energy searching for a more abundant food source. The marginal value theorem (MVT) describes a solution to this dilemma of whether to explore or exploit by stating that

it is optimal to move to a new food source when the instantaneous payoff rate for the resource drops below the average payoff rate in the environment (Charnov et al., 1976). There is evidence that foraging behaviour in a wide range of species, including screaming hairy armadillos (*Chaetophractus vellerosus*; Cassini et al., 1990), guinea pigs (*Cavia porcellus*; Cassini et al., 1990), and starlings (*Sturnus vulgaris*; Vasquez and Kacelnik, 2000), adheres to the MVT.

However, while in general animals might appear to behave optimally, there is variation in foraging behaviour between individuals of the same species (Nonacs, 2001). One factor that has been proposed to underlie this variation around optimality in humans is a person's affective state through its impact on decision-making. The explore/exploit trade-off has most commonly been investigated using variants of the n-armed bandit task in which individuals must choose between a fixed number of options with non-stationary reward values. For example, using this task Blanco et al. (2013) found that humans with depressive symptoms were more exploratory than those without depressive symptoms. However, Blanco et al. (2013) defined exploration as an action which differed from that of an ideal actor, which contrasts with the exploration we describe in this paper. We consider that exploration may be optimal or suboptimal and simply reflects opting to move to new food source as opposed to continuing to harvest the current food source. Exploration on the bandit task in humans has been shown to increase concurrently with depressive anhedonia (Harlé et al., 2017). However, decreased exploration in the bandit task has also been associated with negative affect. For example, stress, as measured by changes in cortisol levels following a stressor and self-reported stress, has been correlated with reduced exploration in humans (Lenow et al., 2017). Yet, with the exception of the explore/exploit task described by Lenow et al. (2017), these n-armed bandit tasks are typically not conducted in a foraging context which would be considered ecologically relevant or to which the MVT would potentially apply. Specifically, consecutive harvests from same food source do not necessarily deplete food availability at that source, which is perhaps unusual in the natural world. Studying exploitation and exploration using tasks which are more ecologically relevant may allow a better understanding the evolutionary significance of affect-induced changes in exploitation and exploration (Scholl and Klein-Flügge, 2018).

Variation in learning and reward sensitivity have been proposed to underlie differences in behaviour exhibited by clinically healthy humans and those with mood disorders (Addicott et al., 2017; Huys et al., 2013; Scholl and Klein-Flügge, 2018). It is well established that depression is associated with a reduced reward sensitivity (Bylsma et al., 2008; Huys et al., 2013). Moreover, anhedonia is one of the diagnostic criterion for depression (American Psychiatric Association, 2013). Computational modelling can provide insights into the cognitive processes underlying exploratory behaviour and allows the influence of learning and reward sensitivity on decision-making to be disentangled. Most commonly, computational models of bandit task data have assumed that exploration is undirected; exploration is random and results from noise

in the decision-process such as the stochasticity generated by a softmax function. In terms of exploratory behaviour, a reduced reward sensitivity may result in an individual making decisions that are more weakly determined by the value of the potential outcomes. This would result in more stochastic decision-making (parameterised as 'inverse temperature' in a softmax function) and consequently greater exploration, particularly in the context where exploitation is optimal (Addicott et al., 2017; Scholl and Klein-Flügge, 2018). Human behavioural studies analysed using computational modelling have found that estimates of parameter characterising reward sensitivity (i.e. inverse temperature) correlate with depressive symptoms (Harlé et al., 2017; Huys et al., 2013), although null results are sometimes observed (Chung et al., 2017; Gradin et al., 2011).

It has also been hypothesised that depression impairs learning, and meta-analyses have concluded that depression is associated with cognitive deficits (Bora et al., 2013; Lee et al., 2012; McDermott and Ebmeier, 2009; Rock et al., 2014). Optimal foraging behaviour relies on the accurate estimation of food availability both at a specific patch and generally in environment which has to be learnt through prior experience. Hence, exploratory behaviour might be influenced by the ability of an individual to learn about the condition of their environment. The extent to which an individual learns from recent rewards or punishers following experience is parameterised as a learning rate. There is conflicting evidence from computational analyses as to whether or not depression is associated with a reduced learning rate. While a number of studies have identified a reduced learning rate in depressed humans (Chase et al., 2010; Dombrovski et al., 2013), several have found no evidence for an affect-dependent learning rate in humans (Dombrovski et al., 2010; Gradin et al., 2011; Huys et al., 2013). An alternative hypothesis is that individuals experiencing negative affective states are less able to modulate their learning rate in accordance with environmental volatility, and there is evidence to suggest that anxious individuals are impaired in their ability to do so (Browning et al., 2015).

The role of affect in altering explore-exploit decisions has received little attention to date in non-human animals. Here we seek to investigate this to establish whether there is evidence that affect manipulations alter decision processes in ways that show parallels to findings in humans. An important additional outcome of this study will be to establish whether decision-making in explore-exploit tasks can thus act as proxy indicators of affect and hence be a useful measure of animal affect with a relatively short training period.

To achieve this, we translated the human bandit task described by Constantino and Daw (2015), in which rewards depleted with multiple visits to the same source, to rats. Given the definition that negative affect is induced by the presence of punishers (anxiety-like states) or absence or removal of rewards (depression-like states; Mendl et al., 2010; Rolls, 2013), enrichment items were temporarily removed from the rats' housing to induce a negative (specifically depressed) affective state. There is a wealth of evidence to suggest that enrichment items are rewarding to rats (Brydges

et al., 2011; Burman et al., 2008a; Patterson-Kane et al., 2001; Van der Harst et al., 2003). The rats completed the bandit task before, after, and during this period of unenrichment. Subsequently, we used computational modelling to characterise behaviour according to three parameters selected because of their relevance in human studies: a learning rate; inverse temperature parameter; and an initial estimate of the average reward rate. We hypothesised that the removal of enrichment would lead to a depression-like state and blunting of reward sensitivity, and result in increased exploration (i.e. a higher inverse temperature parameter); see Fig. 7.1. To further assess this hypothesis, we assessed whether the number of trials completed and volume of juice consumed during the bandit task, as a proxy of motivation to obtain food rewards, differed between test sessions, with the prediction that fewer trials would be completed and lower volumes of juice would be consumed in the unenriched period. Given the conflicting evidence from previous research, we were agnostic as to whether and how the learning rate would be influenced by the affect manipulation.

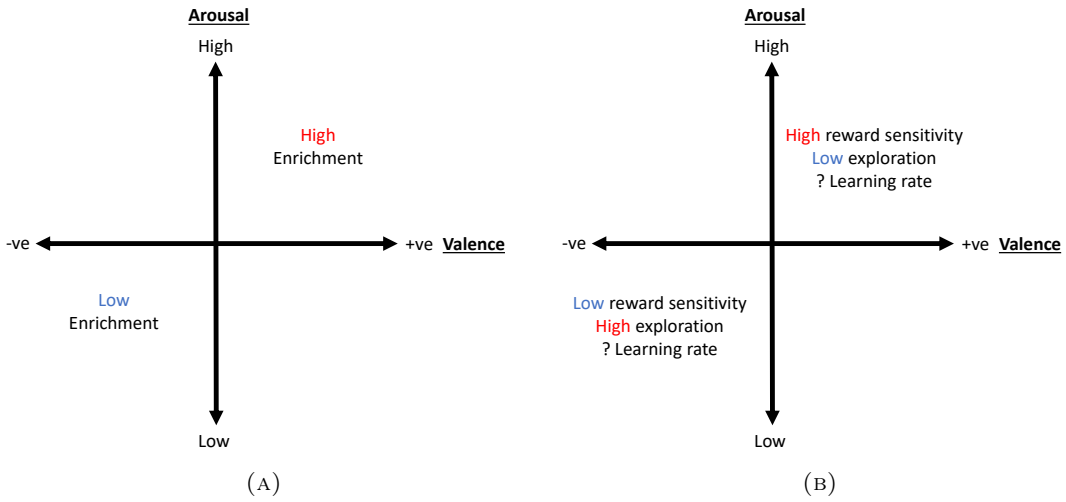


FIGURE 7.1: Diagrammatic summary of predictions: (A) High levels of enrichment would induce a high arousal and positive valenced affective state (excitement-like), while low levels of enrichment would induce a low arousal and negatively valenced state (depression-like); (B) a depression-like state would be associated with blunted reward sensitivity and high exploration.

7.2 Methods

7.2.1 Subjects

Subjects were 20 male Lister Hooded rats (Charles River, Margate, UK), aged c.18 weeks when training commenced. However, only six rats completed testing as a result of irreparable equipment failure and time constraints. The rats had taken part in one behavioural study prior to this experiment. Rats were housed in pairs in cages measuring 560x340x190mm under a 12-hour reversed light-dark cycle (lights on 1900-0700). Cages contained sawdust substrate, shredded paper bedding as well as enrichment as

described below. Food (LabDiet) and water were available *ad libitum*. All rats were checked regularly for any health issues, and the work was conducted under University of Bristol Investigation Number UB/16/004 following approval by the Bristol University Animal Welfare and Ethics Review Body. All rats were rehomed as pets on completion of the study.

7.2.2 Apparatus (see Fig. 7.2)

Training and testing were conducted in a shuttle box (508x254x305mm) placed in a sound isolation chamber. During the first two stages of training, the box was divided in half by a metal panel and one half (254×254×305mm) was used for the training sessions with the other half closed off. The dividing panel was removed during the final training session and testing. A custom-made trough (Faculty of Engineering, University of Bristol) was located centrally on each end wall. An opening (32×40×35mm) which was 35mm above the floor allowed access to the trough from the shuttle box. Syringe pumps controlled the delivery of apple juice to these troughs. A stainless steel tube was inserted through a hole in the side wall of the trough and was connected to the pump-controlled syringe via PVC tubing. All hardware was manufactured by Coulbourn Instruments (Allentown, PA, USA), and operated by their Graphic State (v4) software.

FIGURE 7.2: The shuttle box within a sound isolation chamber used for the explore-exploit task; there are syringe pumps conducted to a trough at either end of the box.

7.2.3 Training and testing

The first two training sessions were 15 minutes in duration and all remaining sessions were 30 minutes in duration. Rats were not food or water restricted prior to training

or testing. In all sessions, rats initiated trials by placing their snout into the trough recess. There were three stages of training. The aim of the first stage was to train rats to associate trough visits with apple juice. This stage comprised two training sessions, each conducted in separate sides of the shuttle box, in which 0.1ml of apple juice was delivered 20ms after the rat's snout entered the trough. The aim of the second stage, which comprised a minimum of two sessions and was also conducted in separate sides of the shuttle box, was to train the rat to keep his snout in the trough for one second to obtain 1ml of juice. To progress to the final training stage, rats had to complete at least 60 trials and stay in the trough for 1s on significantly more trials than chance, as determined by a binomial test, for two consecutive training sessions. The aim of the final training stage was to introduce the rat to the bandit task, which would involve visiting both troughs. Sessions in the final training stage had the same format as the test sessions, and to progress to testing rats had to complete at least 20 trials with at least 5 visits to each trough. Throughout the final training stage and during test sessions both troughs were available. On the initial visit to each trough, and on visits where the rat had previously been at the alternate trough, the volume of apple juice delivered was drawn from a Gaussian distribution with a mean of 0.1ml and a standard deviation of 0.005ml. The volume of juice delivered reduced on consecutive visits to the same trough, with the juice reduction determined by a Gaussian distribution with a mean of -0.01ml and a standard deviation of 0.005ml (excluding values greater than 0). If the rat did not stay in the trough for 1s then no apple juice was delivered but the juice reward continued to deplete. Following the tenth consecutive visit to a trough, the volume of juice delivered was determined by a Gaussian distribution with a mean of 0ml and a standard deviation of 0.005ml which was bounded at zero. Therefore, in the task rats had to make a choice to stay at the same trough with a depleting juice supply ('harvest') or to forgo juice to move to the other trough ('switch'). There were three test sessions in total: the first test session was conducted the day following the final training session, the second test session was conducted three days following the removal of enrichment, enrichment items were returned immediately following the second test session, and the third test session was conducted three days following the return of the enrichment items. Rats had a practice test session on the day prior to the unenriched and enrichment-return test session to refamiliarise them with the task.

7.2.4 Environmental manipulation

Unenriched housing has been associated with poor welfare and negative affect in rats (Brydges et al., 2011; Burman et al., 2008a; Van der Harst et al., 2003), and rats show a preference for enriched cages (Patterson-Kane et al., 2001). Hence, we manipulated housing conditions to induce changes in affect. Rats were provided with enrichment from their time of arrival at the University of Bristol. This comprised an aspen block, a small cardboard tube, a plastic tube attached to the cage ceiling, and a cardboard nest box (Fig. 7.3). With the exception of the cardboard tube, all enrichment was

removed immediately following the first test session (Fig. 7.3), and the enrichment was returned immediately following the unenriched test session.

(A) (B)

FIGURE 7.3: Examples of (A) enriched and; (B) unenriched housing

7.2.5 Statistical analysis

The key outcome variables were the decision to 'harvest' or 'switch' and the number of trials completed in each session. A generalised linear mixed model (GLMM), using the R (R Core Team, 2017) packages nlme (Pinheiro et al., 2018) and lme4 (Bates et al., 2015), which included a random effect of session nested within individual, was fitted to the binary decision to 'harvest' or 'switch' with housing type (enriched vs. unenriched), number of test sessions completed, and volume of juice delivered on the previous trial as predictor variables. A generalised linear mixed model with a Gaussian error structure and random effect of individual was fitted to the number of trials completed in each session. Housing type and number of test sessions completed (to control for potential time or experience dependent effects on behaviour) were predictor variables in this model. To examine whether the treatment and number of sessions completed influenced whether rats were willing to wait 1s for the apple juice, a Gaussian GLMM was also fitted to the proportion of successful harvests in each test sessions with housing type and number of sessions completed as predictor variables. The model assumptions were verified and likelihood ratio tests (LRT) were used to assess whether the difference in model deviance was significant when a predictor variable was removed from the model.

7.2.6 Marginal value theorem model

Following Charnov et al. (1976) and Constantino and Daw (2015), we consider that it is optimal for an individual to 'switch' when the expected reward from an additional visit is less than the opportunity cost of the time spent harvesting the reward. This rule for switching can be formalised as:

$$\pi^* = \begin{cases} switch, & \text{if } \epsilon_{t+1} - h\bar{R}_t < 0 \\ harvest, & \text{otherwise} \end{cases} \tag{7.1}$$

where ϵ_{t+1} is the expected reward, $\bar{R}_t$ is the average reward rate at time t, and h is the time taken to harvest the reward at the current trough (1s in this study). The average reward rate incorporates the time taken to move between food sources; this was demonstrated by Constantino and Daw (2015). More specifically, the average reward rate decreases as the time taken to move between food sources increases (resulting from the increased time spent travelling instead of foraging).

Both ϵ_{t+1} and $\bar{R}_t$ can be defined using a Rescorla-Wagner learning rule (Rescorla et al., 1972). Specifically, ϵ_{t+1} reflects the outcome of the previous harvests $r_{0:t}$ and updates according to learning rate α on consecutive successful visits to the same trough:

$$\epsilon_{t+1} = \epsilon_t + \alpha(r_t - \epsilon_t) \tag{7.2}$$

The value of ϵ_{t+1} following a trough switch was the outcome of the harvest r_t. We verified the assumption that the expected reward was not updated following unsuccessful harvests and also that rats learned the expected reward from experience as opposed to updating based on the true depletion rate (i.e. depletion by mean=-0.01ml) by fitting models using these alternate updating strategies and comparing the model fit. The AIC value indicated a poorer fit in each case (fitting all models including unsuccessful harvests: ΔAIC=77.489; fitting all models with update using true depletion rate: ΔAIC=369.056, in comparison to all models not including unsuccessful harvests and no updating with the true depletion rate).

$\bar{R}_t$ is given initial value $\bar{R}_0$ and updates following each harvest according to the outcome of and the time since the last harvest τ:

$$\bar{R}_{t+1} = \bar{R}_t + (1 - (1 - \alpha)^\tau)(\frac{r_t}{\tau} - \bar{R}_t) \tag{7.3}$$

We fitted an additional model in which the average reward rate and expected reward update had different learning rates but this was found to provide a less parsimonious model (ΔAIC=1.574, comparing all models using separate learning rates for the average reward rate and expected reward, to the same set of models using the same learning rate for the average reward rate and expected reward update).

The probability that an individual continues to harvest at the same trough is determined using a softmax function with inverse temperature β (Fig. 7.4; Table 7.1):

$$P(\text{harvest}) = \frac{1}{1 + e^{(-\beta(\epsilon_{t+1} - \bar{R}_t))}} \tag{7.4}$$

The softmax function is widely used in the field of reinforcement learning and outputs an action probability given the action values and the inverse temperature

parameter (Sutton and Barto, 1998).

Table 7.1: Glossary of MVT model parameters

Parameter	Range	Interpretation
α	$[0, 1]$	Learning rate for the average reward rate and expected reward; higher values reflect faster updating of the average reward rate and expected reward following decision outcomes.
β	$[0, \infty]$	Inverse temperature parameter; higher values reflect that decisions are less stochastic and more heavily based on the expected value of actions.
R_0	$[0, \infty]$	Initial value of the average reward rate; higher values reflect that the average reward rate is higher at the start of the test session.

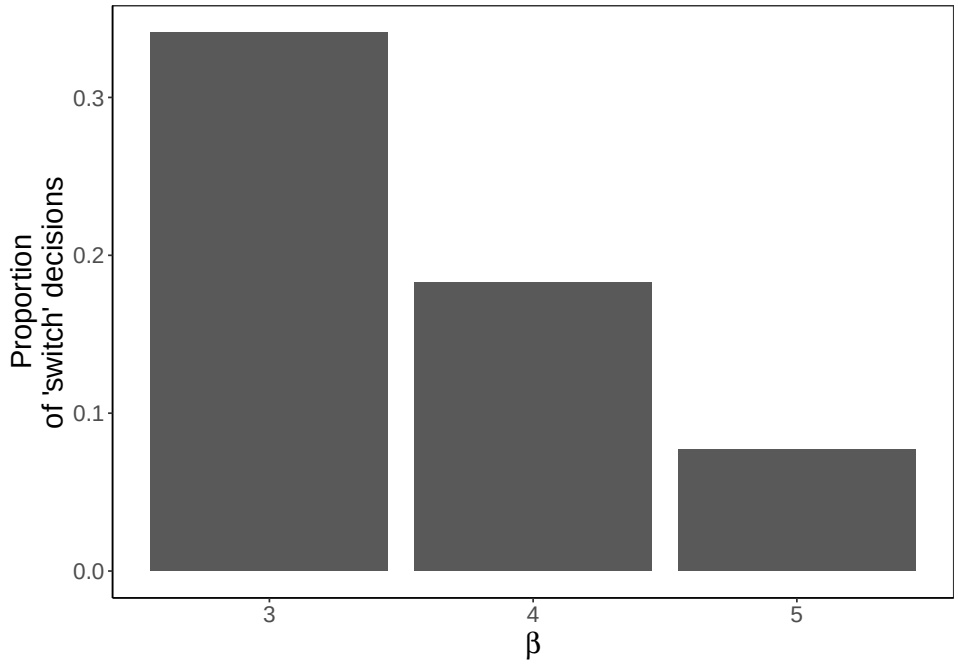


Figure 7.4: Data generated using the MVT model to demonstrate β-dependent variation in decision-making

We used a three-level hierarchical Bayesian random effects analysis for model fitting which was previously described in Chapter 5. Briefly, this involved parameterising distributions for the estimates of α, β, $\bar{R}_0$ from each test session for each individual, and parameterising the estimates of these individual-level distributions as a higher-level distribution across all subjects (as described in Chapter 5). The top-level parameters were estimated using an expectation-maximisation procedure (Guitart-Masip et al., 2012; MacKay and Mac Kay, 2003).

The likelihood function and code for the EM procedure were written using MATLAB and each model took less than 10 minutes to fit (and hence did not required use of the high-powered computing facility). The initial values for all parameters were a number drawn at random between zero and one. Values for parameters constrained

between zero and ∞ were exponentiated within the likelihood function, likewise, values for parameters constrained between zero and one were transformed using a logistic function within the likelihood function.

7.3 Results

All six rats completed training in five days. The rats completed an average of 28.250±6.868 (mean±SE) trials in the first training stage in which they kept their snout in the trough for 20ms (i.e received juice) on 98.846±0.008% of trials, an average of 47.833±6.498 trials in the second training stage in which they in which they kept their snout in the trough for 1s (i.e received juice) on 77.426±0.045% of trials, and an average of 57.833±12.202 trials in which they kept their snout in the trough 1s (i.e received juice) on 55.062±0.029% of trials during the final training session. This decline in performance perhaps reflects the reduction in reward magnitude resulting from repeated trough visits in the final test session.

Across test sessions, rats completed an average of 95.278±6.808 trials and in which they kept their snout in the trough to successfully harvest the juice on 62.492±0.024% of trials. Rats completed significantly fewer trials in the unenriched condition (Fig. 7.5: LRT=8.234, p=0.004) and the number of sessions completed was not a significant predictor of number of trials completed in a session (LRT=2.337, p=0.126). However, the volume of juice consumed did significantly depend on the number of sessions completed, with juice consumption increasing with the number of sessions (LRT=11.218, p<0.001). Total juice consumption did not depend on housing type (LRT=0.005, p=0.944). The proportion of successful harvests did not depend on housing type (LRT=2.647, p=0.104) or number of sessions completed (LRT=0.514, p=0.473).

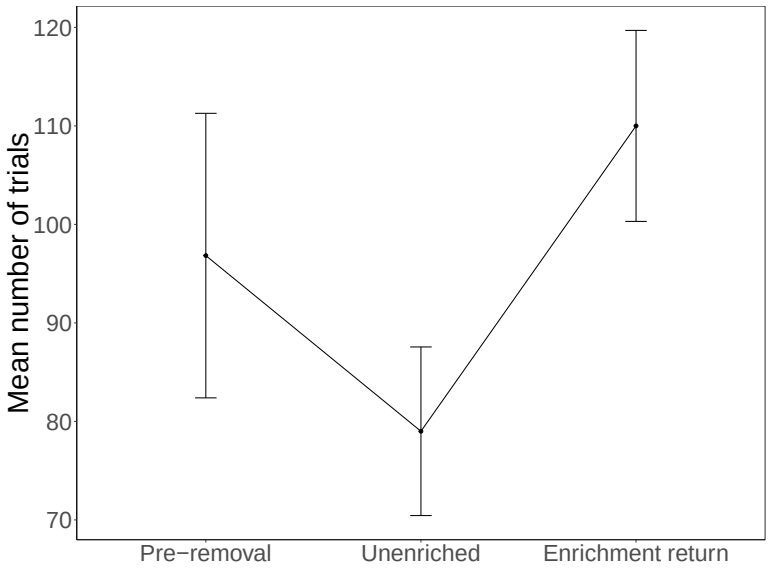


FIGURE 7.5: The mean number of trials completed in each test session. Error bars represent one standard error

Rats were more likely to 'switch' in the unenriched condition (Fig. 7.6: LRT=5.482, p=0.019). The number of test sessions completed was not a significant predictor of the decision to 'switch' or 'harvest' (LRT=0.360, p=0.548).

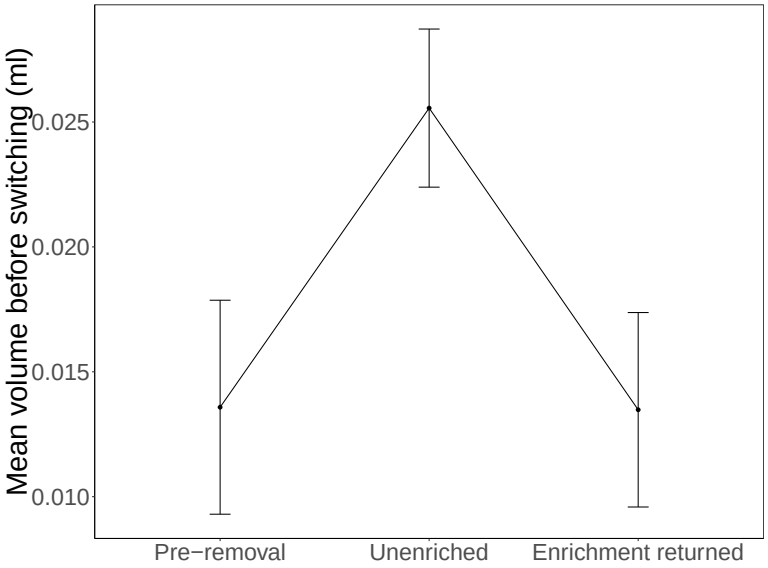


FIGURE 7.6: The mean volume of juice harvested at a trough before the decision to switch to the alternate trough in each test session. Error bars represent one standard error

Housing type was a significant predictor of both α (learning rate; Fig. 7.7: LRT=116.345, p<0.001) and β (inverse temperature; Fig. 7.8: LRT=4.886, p=0.027) but not R_0 (initial estimate of the average reward rate; LRT=0.123, p=0.726). The parameter estimates of both α and β were lower when rats were housed in unenriched cages. There was no significant effect of number of test sessions completed on α (LRT=1.452, p=0.228) and β (LRT=0.970, p=0.325) or R_0 (LRT=1.675, p=0.196).

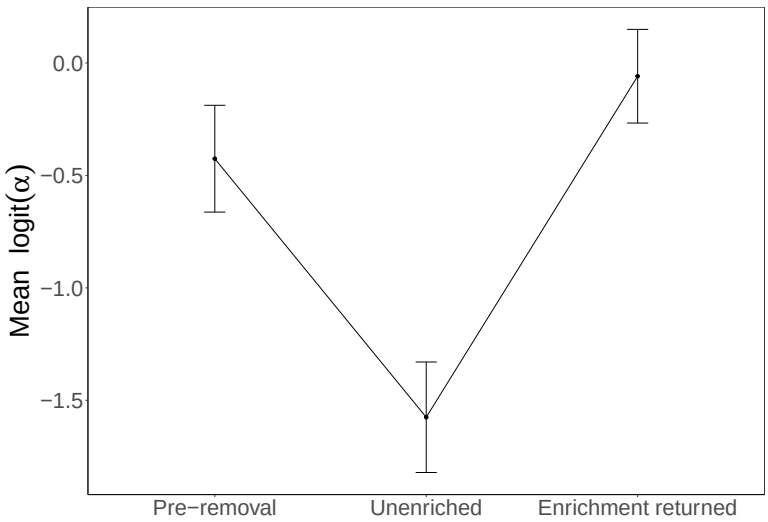


FIGURE 7.7: The mean parameter estimates of α in each test session. Error bars represent one standard error

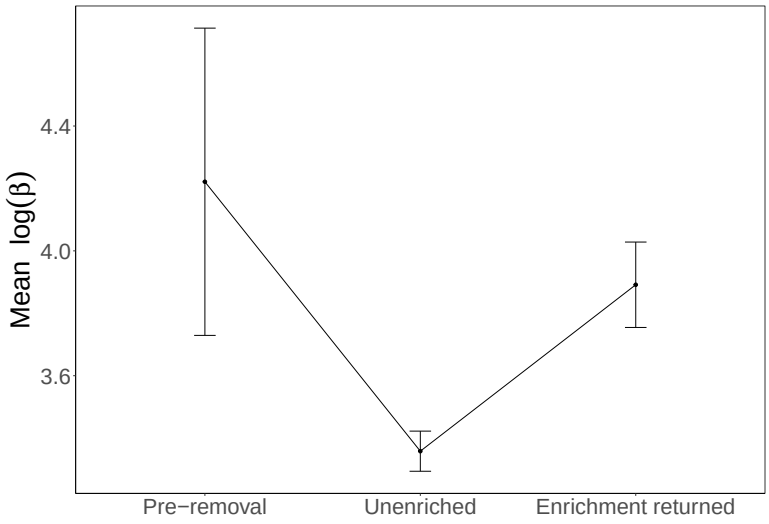


FIGURE 7.8: The mean parameter estimates of β in each test session. Error bars represent one standard error

7.4 Discussion

The aims of this study were to use computational modelling to examine the specific cognitive changes contributing to affect-induced shifts in exploration on the bandit task and to assess whether bandit tasks could provide a measure of affective state in rats. To this end, we developed a rat bandit task in which successive visits to the same trough led to reduced rewards which was adapted from the human task described by Constantino and Daw (2015). Data from this task were analysed using a computational model based on the marginal value theorem.

Rats in putatively negative affective states resulting from less-enriched housing conditions were more exploratory, which was attributed to a lower inverse temperature parameter. Specifically, according to the model-based analysis their decision-making was less dependent on the relative value of 'harvesting' compared with 'switching'. This is consistent with research from the field of computational psychiatry that has identified lower estimates for the inverse temperature parameter in data obtained from depressed humans (Harlé et al., 2017; Huys et al., 2013), which may reflect a blunted reward sensitivity.

The affect manipulation also influenced learning; rats had lower learning rates when in a putatively negative affective states. This concurs with the hypothesis that depression is associated with impaired learning (Chase et al., 2010; Dombrovski et al., 2013) and suggests that rats in depression-like states might also exhibit learning impairments. However, it is important to note that it is unclear what an altered learning rate might reflect in terms of cognition; although it could reflect differences in learning, it could alternatively reflect differences in the individual's ability to encode or recall reward experience, differences in perception of the magnitude or value of the

experienced reward, or reflect differences in the individual's estimate of environmental volatility given that volatility is a key determinant of learning rate (Browning et al., 2015). Disentangling these possibilities will be important to better interpreting the learning rate in studies of human depression and animal welfare. Overall, the behaviour and cognition of rats from the less-enriched housing condition appears to closely mirror that of depressed humans.

Rats also completed fewer trials in test sessions when their housing was unenriched, which might be indicative of a reduced motivation to obtain rewards. This also needs to be interpreted in the context of total volume of juice consumed which increased across the test sessions. Taken together, it suggests that although the number of trials decreased in the unenriched session, the absolute juice consumption did not decrease, and so the greater exploratory behaviour resulted in more efficient foraging (in terms of juice per harvest) relative to the test session prior to enrichment removal. If we consider increased juice intake as a measure of optimality, this also suggests that a reduced learning rate and inverse temperature parameter are not necessarily associated with suboptimal decision-making. This contrasts with the hypothesis that depression leads to suboptimal behaviour (Montague et al., 2012; Paulus and Angela, 2012). However, given that juice intake could be increased with faster shuttling and so also a greater energy expenditure, interpretation of the results in terms of optimality is complicated by a lack of measure of the energetic cost of harvesting and shuttling between troughs. The increased juice consumption across test sessions might also indicate that increased familiarity with the task improves performance.

Although the aim of this study was not to address different potential exploration strategies it is important to briefly discuss directed and undirected exploration (Carmel and Markovitch, 1990; Sutton, 1990; Thrun, 1992). The MVT and the majority of computational models used to analyse exploration and exploitation data in the field of computational psychiatry have assumed that exploration is undirected; the decision to explore or exploit depends on the utility of each option with some stochasticity. However, in some contexts exploration can be used to reduce uncertainty, and hence exploration may be valuable to better understand environmental conditions. Directed exploration describes a strategy which accounts for the added value of exploration. In future studies it may be worthwhile to assess whether rats might also adopt a directed learning strategy and to examine whether affect might influence the extent to which exploration is directed or undirected, and hence underlie differences in exploratory behaviour between individuals in putatively negative and positive affective states.

Although our results suggest that the bandit task is a promising measure of affective state in rats, which requires little training, it is important to highlight a few caveats. Firstly, the small sample size in this study may reduce the reliability of the results (Button et al., 2013). Secondly, while there is strong evidence to suggest that removal of enrichment should induce a relatively negative affective state, it is

nonetheless an assumption. Finally, as the putative affective state induced by the manipulation most closely resembles a depression-like state, it is unclear how successful the task would be in detecting anxiety-like states. Consequently, the task would benefit from further investigation with larger sample sizes. Pharmacological validation of the task would also be important to determine its value as a measure of affect.

From a welfare perspective, a crucial point is that the less enriched housing condition in our study may be marginally more preferable than typical housing for laboratory rats. In the less enriched housing condition, rats had more floor space than current regulations require and rats were still provided with one form of enrichment (UK Home Office, 2014). Our results, and indeed that of other researchers (Brydges et al., 2011; Burman et al., 2008a; Van der Harst et al., 2003), might indicate that changes in housing would enhance the welfare of laboratory rats, which would ultimately improve the scientific value of experimental studies that use rats (Poole, 1997; Prescott and Lidster, 2017; Sherwin, 2004).

7.5 Conclusions

The bandit task we have described may provide a useful measure of affective state in rats which can be analysed through computational modelling. Rats in putatively negative affective states were more exploratory and completed fewer trials in this bandit task. Additionally, the computational analysis revealed that less-enriched housing conditions increased stochasticity in decision-making and reduced the individual's learning rate, indicative of a depression-like state. The task therefore warrants further investigation as a measure of affect.

Chapter 8

General Discussion

The judgement bias task has been widely used to assess affect, and hence welfare, in a range of non-human animals. However, a solid understanding of how and why affect alters judgement bias is still lacking. As outlined in the general introduction (Chapter 1), it has been hypothesised that affect might encode different aspects of reward and punisher experience, such as predictability or how relatively or absolutely appetitive or aversive an individual's experience has been, to provide a context for decision-making, hence allowing appropriate behavioural responses. Reward and punisher experience could directly, or indirectly via affect, alter decision-making by influencing subjective reward and punisher valuation or probability estimation. This hypothesis is supported by research in humans (see Chapter 1). To this end, the aim of this thesis was to examine the relationship between affect, reward and punisher experience, and the cognitive processes underlying decision-making, and also to look for parallels/differences between humans and rats in order to identify cross-species (dis)continuity in links between environmental conditions, affect, and decision-making. In this general discussion we will assess the extent to which this aim was achieved, and our hypotheses supported, as well as discussing the implications and limitations of our findings and providing suggestions for future research. To avoid repetition, as results are discussed in each experimental chapter, we will only provide an overview of the results in this discussion where context is required.

8.1 Which aspects of reward and punisher experience influence decision-making?

Both humans and rats were more likely to risk a punisher to obtain a reward when the expected value of the trial was greater (Chapters 3, 4, and 6). This suggests that both humans and rats attend to the potential outcome of each option and use this information to inform and optimise decision-making. Hence, this indicates that individuals make decisions in accordance with theories of decision-making (see Chapter 1) and provides validation for the task as a measure of decision-making.

Humans were more risk-averse when recent outcomes had been less predictable/-more surprising both when primary and secondary rewards and punishers were offered or threatened. This could reflect a strategy to reduce uncertainty, although opting

for the risk-averse response, whilst reducing outcome uncertainty, also reduces the opportunity to learn and hence minimise future uncertainty (Chapters 3 and 4). In contrast, the predictability of recent outcomes did not influence decision-making in rats, but instead rats (although only one individual rat significantly) were more risk-averse when outcomes had been absolutely less favourable (Chapter 6). This would allow rats to make decisions in accordance with environmental conditions.

Consequently, we have identified disparities in the relationship between experience and decision-making in both humans and rats. This inconsistency could be due to differences in the focus of the individuals when completing the task; resolving uncertainty may be the most important goal to humans, while maximising rewards and minimising punishers may be more important to rats. Laboratory rats inhabit a predictable and controlled environment; they live in stable groups, experience negligible fluctuations in ambient temperature and humidity, and husbandry and testing occur at regular and predictable times. In contrast, humans regularly interact with different people, experience fluctuating environmental conditions, and may have a varied daily schedule. This may induce a greater drive in humans to reduce uncertainty (at least of immediate outcomes) compared with laboratory rats. Additionally, the rewards and punishers used may be more salient to rats than to humans. The sucrose pellet may be especially rewarding to rats who have a diet consisting only of standard laboratory chow, particularly given that wild rats are opportunistic and omnivorous feeders (Schein and Orgain, 1953). Moreover, as a prey species, the potential punisher may be considered life-threatening to rats. Indeed, in our experiment, equipment failure that resulted in an air-puff not stopping resulted in a rat displaying tonic immobility indicative of perceived imminent death and hence an extreme fear-like state[1]. This may promote decision-making that is informed by the absolute favourability of rewarding and punishing experiences. On the other hand, there may be far more appealing foods to humans, who experience a range of food daily, than apple juice, and the monetary gains from the task were minimal, especially using the UK minimum wage (currently £6.15 for those aged 18-20; £7.70 for those aged 21-24; and £8.21 for those aged over 24) as a reference point. Additionally, the humans were made aware that the experiment received ethical approval and that the experiment posed no risk to their life.

Furthermore, humans completed the task with a conspecific present i.e. under observation by the experimenter. Thus, there may be social factors influencing human decision-making. For example, humans may want to perform well to satisfy the experimenter. Consequently, the increased predictability of recent outcomes during testing, taken as evidence of good comprehension of the task, might inspire greater confidence and less cautious decision-making in humans.

The prediction error most relevant to decision-making also differed between humans and rats. While the results of the computational analysis suggested that for

[1]This rat was immediately withdrawn from the experiment and observed until he had made a full recovery. It should also be noted here that some rats do habituate to, and subsequently show no avoidance of, air-puffs

humans it was the difference between the outcome and prediction about the outcome prior to stimulus onset that was most relevant, the difference between the outcome and prediction about the outcome following stimulus onset was most relevant to rats. This difference potentially stems from an important methodological difference; humans were shown the potential outcomes prior to stimulus presentation whereas the specific potential outcome on each trial was unknown to rats, and their pre-stimulus onset prediction would have been based on experience. It may be that the greater degree of certainty about the potential reward/loss magnitude or simply being forced to consider the potential outcomes prior to stimulus presentation made the pre-stimulus prediction more pertinent to humans compared to rats.

Both humans, where primary as opposed to secondary rewards and punishers were used (Chapter 4), and rats when within-test experience was manipulated (Chapter 6) initiated trials with greater alacrity when recent outcomes had been more predictable. We interpreted this to reflect a means to process and learn from unexpected outcomes, or to search for novel cues to provide information about changing environmental conditions. Hence, there is evidence that rats encode information about predictability, even if predictability does not modulate their decision-making.

However, predictability did not exert the same effect on vigour in humans when secondary rewards were used (Chapter 3), or on vigour in rats where reward and punisher experience was manipulated prior to testing (Chapter 5). Instead, humans in the monetary task were slower to initiate trials following outcomes that were better than expected (consistent with slowing following decreased predictability), but faster following outcomes that were worse than expected (inconsistent with slowing following decreased predictability). We suggested that this may reflect differences between punishers and losses; losses can be recouped but punishers cannot be undone, hence leading participants to increase vigour following unexpectedly poor monetary outcomes so as to more rapidly regain attempts to increase earnings in line with expectation. In rats, when pre-test reward and punisher experience was manipulated, variability in trial initiation latency could be attributed to the outcome of the previous trial. This could reflect that the within-test reward magnitude was static so that predictability of recent outcomes is less relevant to shifts in environmental conditions.

The average reward or average earning rate influenced vigour in both rats (Chapter 6) and humans (Chapters 3 and 4). In particular, individuals were faster to initiate trials when they had learnt that the environment was more favourable. This is consistent with both theoretical and empirical studies investigating the relationship between reward experience and vigour (Beierholm et al., 2013; Griffiths and Beierholm, 2017; Niv et al., 2007), and is thought to optimise the trade-off between energetic expenditure and reward intake given the environmental conditions.

Therefore, a key finding of this thesis is that reward and punisher experience modulates decision-making in the judgement bias task and vigour in rats and humans. While there are some similarities in the specific aspects of reward and punisher experience which influence behaviour in rats and humans, there are also differences

which require further examination.

8.2 Which cognitive processes underlie the relationship between reward and punisher experience and decision-making?

In the human monetary rewards and losses study, participants overall had a reduced sensitivity to losses relative to the true monetary value of the loss (Chapter 3), whereas humans in the primary reward and punisher task (Chapter 4) and rats (Chapter 6) showed an increased sensitivity to rewards relative to the true volume of apple juice offered/received. One possible explanation for this is that the losses might not be perceived as true losses in a monetary task given that it is not possible for humans to finish the task with less money than they had prior to testing, and participants were aware that the experimenter cannot take money from them. On the other hand, in the primary reward and punisher task, participants were asked to abstain from drinking (other than water) and eating in the hour prior to testing, which through hunger may have inflated the subjective value of the apple juice. Although rats were not food-deprived for any period of time during experimentation, the judgement bias training and test sessions were the rats' sole opportunity to consume food other than laboratory chow and hence may have been highly rewarding.

It is also important to note that only the reward magnitude fluctuated across testing in the human primary reward and punisher task, while there was both a fluctuating reward and fluctuating loss condition in the human secondary reward and punisher task. Although, this could potentially underlie the differences in the cognitive processes underlying judgement bias between studies, in the monetary task participants were less sensitive to loss in the fluctuating reward condition in comparison to the fluctuating loss condition which suggested that the finding would still be observed if only a fluctuating reward condition had been used.

Variation in the prior probability of reward also governed judgement bias in the human primary reward task. Participants overall had a greater expectation that the trial would be punished than rewarded. This could reflect that participants had preconceptions about what the task would involve, leading them to believe that delivery of salty tea would be more probable than delivery of apple juice. Given that an increased subjective probability of punishers has been associated with negative affect, this finding could also reflect that overall the humans in the primary reward task were experiencing a relatively negative affective state compared with participants in the monetary task. This negative affect could arise from apprehension about the experiment, for example concerns about an embarrassing reaction to the salty tea (e.g. coughing, dribbling) or because the ambient room temperature was noticeably cool as only a portable heater was available to heat the experimental room (with the study conducted in Winter in a non-heated building). Room temperature has been shown

to influence affective state in humans (Bell and Baron, 1977) and chickens (Deakin et al., 2016). Although ideally this explanation would be examined by comparing the results of the affect grid between the primary and secondary reward and punisher tasks, this was unfortunately not possible as testing in each study was conducted using a different monitor with different number of pixels and so the co-ordinates are not comparable and the absolute location within the grid might also not be comparable due to differences in the speed at which each position within the grid could be reached.

The computational analysis revealed that the subjective valuation of rewards or punishers was influenced by reward and punisher experience, which contributed to within-subject variation in judgement bias in both rats, where the absolute favourability of experience influenced reward valuation (Chapter 6), and humans where the predictability of recent outcomes influenced loss sensitivity (Chapters 3) or reward sensitivity (Chapter 4). This was further supported by a concurrent measure of motivation to obtain rewards described in Chapter 5; rats that experienced pre-test sucrose valued punishers more highly than rewards and made fewer presses to obtain sucrose in the progressive ratio lever pressing task compared to rats that experienced pre-test air-puffs. This finding firstly indicates that judgement bias, which is largely considered to reflect affect-induced changes in probability estimation (Bateson et al., 2011; Mendl et al., 2010; Nettle and Bateson, 2012) could also, or instead, arise from variation in the subjective valuation of rewards and punishers. Moreover, as highlighted in Chapter 1, the influence that affect may exert on the subjective valuation of rewards and punishers is unclear. Thus, a negative or positive judgement bias might not necessarily be indicative of the relative valence of an affective state. It is possible that a relatively negative judgement bias could reflect a reduced valuation of rewards or increased valuation of punishers resulting from a relatively positive affective state, while it is also possible that a relatively positive judgement bias could reflect an increased valuation of rewards or decreased valuation of punishers resulting from a relatively negative affective state. Indeed, this could explain the findings of Chapter 3 where delivery of sucrose to rats, assumed to induce a relatively positive affective state relative to delivery of punishers, induced a negative judgement bias, and also provide an explanation for other opposite findings in the literature (e.g. Briefer and McElligott, 2013), as well as the high levels of heterogeneity across judgement bias tasks revealed by the meta-analysis (Chapter 2).

However, in the human primary reward and punisher study (Chapter 3), reward and punisher experience (specifically the predictability of recent outcomes) also influenced the participants' prior expectation that the trial would be rewarded or punished. Thus, reward and punisher experience does not solely modulate reward or punisher valuation. Therefore, the question remains why both subjective reward and punisher valuation and probability estimation were modulated by experience in the human primary reward and punisher study but not in other judgement bias studies described in this thesis. A possible explanation is that there was greater power to detect this effect in the human primary reward and punisher study; the sample size in the rat

task was unfortunately very small due to problems with training, and given that the immediate rewards are potentially more engaging to humans than promised monetary rewards, the effect size may have been comparatively large.

8.3 The relationship between reward and punisher experience and decision-making and affect

The meta-analysis (Chapter 2) generally supported the hypothesised relationship between affect and judgement bias, suggesting that relatively negative affect is associated with a negative judgement bias, and that a relatively positive affect is associated with a positive judgement bias, across a range of species. However, there was wide heterogeneity in the extent to which the pharmacological manipulations altered judgement bias which suggested that while some studies had small effect sizes or altered judgement bias in the opposite direction to expected other studies had large effect sizes. Thus, while in general judgement bias may provide a measure of affect, caution must be taken to interpret judgement bias data.

With regards to reward and punisher experience, although recent unpredictability in outcomes was found to influence judgement bias in humans, it did not influence subjectively reported affective valence. However recent unpredictability in outcomes did influence affective arousal; greater arousal was reported when outcomes were less predictable. Conversely, the absolute favourability of experience determined affective valence in humans, with more positive affect reported when experience had been more absolutely favourable, without influencing judgement bias. This highlights that the aspects of reward and punisher experience which influence affect do not necessarily directly alter judgement bias, while the aspects of reward and punisher experience which influence judgement bias do not necessarily directly alter affect. Consequently, it appears that reward and punisher experience are not necessarily encoded by affect to influence decision-making. This finding also indicates that judgement bias may be related to affective arousal which is less relevant to animal welfare.

However, affect indirectly influenced decision-making by modulating the extent to which unpredictability informed judgement bias. More specifically, in both Chapters 3 and 4, which detail studies using humans, reports of more negative affective valence were associated with unpredictability exerting a weaker influence on the subjective valuation of the rewards or losses. This could reflect that negative affect is associated with blunting of prediction error signalling (Delgado et al., 2011; Kumar et al., 2008; Steele et al., 2004), or could reflect an effort to maintain a positively valenced affective state (through increased risk-aversion) in periods of uncertainty. Additionally, in the human primary reward and punisher task (Chapter 4), participants that reported more negative affective valence exhibited an attenuated belief that the trial would be rewarded in periods of greater unpredictability. This could indicate that the effects of affective valence on decision-making are greatest when the environment is perceived as unpredictable. Consequently, affect might not only determine judgement bias (as

demonstrated in Chapter 2) but also determine within-subject variability in judgement bias, and there may be interactions between environmental conditions and the influence of affect on judgement bias, such that the influence of affect on judgement bias might only be observed when recent rewarding and punishing outcomes have been less predictable.

In contrast to previous research (i.e. Rutledge et al., 2014), the relative favourability of the environment (prediction error) did not influence affective valence. Instead, the most recent outcome provided a better account of within-subject variation in affect, with humans reporting more positive affective valence when the most recent outcome was more favourable (Chapters 3 and 4). Likewise, humans reported more positive affective valence when the expected value of the most recent trial was greater, as well as greater arousal when the magnitude of the expected value was greater. This provides support for Roll's (2013) definition of emotion as states elicited by rewards and punishers.

While the experimental chapters in this thesis support elements of current theories of affect, they do not wholly support any particular theory. Firstly, we did not find that more predictable environments were associated with more positively valenced affect, as suggested by both Huys and Dayan (2009) and Clark et al. (2018). However, we did find that affective valence modulated the extent and direction in which unpredictability modulated decision-making. The finding that participants in more negatively valenced affective states exhibit decision-making that is less dependent on the predictability of recent outcomes could be considered indicative of a less flexible decision-making strategy. This would not be contrary to the hypothesis that affect determines the strength of the belief that the world is predictable, where a strong belief (high precision around prior beliefs) that the world is unpredictable, theorised to be associated with depression (Clark et al., 2018), would lead to weaker reliance on environmental information. Notably, Browning et al. (2015) found that humans were unable to adjust their learning rate in accordance with environmental conditions, and it has also been demonstrated that negative affect induces a shift towards model-free (which is less flexible that model-based) decision-making (Blanco et al., 2013; Huys et al., 2012; Radenbach et al., 2015). Thus, in combination these findings are consistent with an association between affect and behavioural flexibility. It would be worthwhile in future studies to more directly assess the hypothesis that negative affect is associated with model-free decision-making strategies, for example using tasks designed for rats (e.g. Hasz and Redish, 2018) and humans (e.g. Daw et al., 2011b) that in combination with computational modelling, allow use of model-based and model-free decision-making to be disentangled, alongside an affect manipulation.

Contrary to the hypothesis that affect reflects the relative favourability of environmental conditions (Eldar et al., 2016), we did not find that affect reflects the difference between observed and expected outcomes (prediction error). However, the finding that the absolute favourability (average earning rate) influenced human affect is consistent with the hypothesis that affective valence reflects an individual's

cumulative experience of rewards and punishers (Mendl et al., 2010). Although, in contrast to the hypothesis of Mendl et al. (2010), the average earning rate did not influence human decision-making, and so there is no evidence to suggest that affect functions to provide information about the absolute favourability of the environment to inform human decision-making. However, the average earning rate influenced rat judgement bias suggesting that the average earning rate may function to inform rat decision-making within the judgement bias task.

In both rats and humans, the absolute favourability of the environment modulated vigour, suggesting that affect (as a measure of environmental favourability) could function to determine the vigour with which actions are executed. Moreover, serotonin and dopamine have been proposed to underlie both vigour and affect, hence providing a potential neurobiological explanation for this finding (Berridge and Robinson, 1998; Boureau and Dayan, 2011; Niv et al., 2007; Owens and Nemeroff, 1994). However, we would expect vigour and arousal to be closely related. Thus, this finding raises the question of why the absolute favourability of the environment determines affective valence but not affective arousal in human studies using the affect grid, given that it also determines vigour. It could suggest that participants have difficulty separating valence and arousal when reporting affect using the affect grid, although expected value dependent shifts in affective arousal were successfully detected using the affect grid.

To summarise, affective valence in humans might function to modulate the extent to which information about unpredictability, obtained from reward and punisher experience, modulates decision-making. Adopting a less flexible decision-making strategy could reduce energy expenditure by decreasing cognitive load. The absolute favourability of previous outcomes modulates the vigour with which actions are executed in both rats and humans, and it also determines decision-making in rats. In humans, the absolute favourability of previous outcomes also determined affective valence and so it is plausible that the effect of absolute environmental favourability on vigour (in both humans and rats) and decision-making (in rats) operates via affective valence. Further studies with mediation analyses which assess the extent to which one variable influences another variable directly, or indirectly via a mediating variable would be necessary to confirm this. Modulating vigour and decision-making in accordance with the absolute levels of rewards and punishers in the environment would allow reduced energy expenditure and more cautious decision-making in less favourable environments, and increased energy expenditure and less cautious decision-making in more favourable environment. Hence, overall our research suggests that affect may function to mediate energy expenditure, allowing energy to be conserved in poorer environments. It would be worthwhile to directly test this hypothesis by taking direct measures of energy expenditure such as oxygen consumption (e.g. Jequier et al., 1987).

8.4 Implications for judgement bias as a measure of welfare

Our findings have clear implications for judgement bias as a measure of welfare. Firstly, although overall judgement bias appears to measure affective valence (as per the meta-analysis, Chapter 2), the high level heterogeneity identified in addition to the finding that judgement bias may reflect reward and punisher valuation (Chapters 3 - 6) signifies that caution is necessary when interpreting the results of judgement bias. It is not necessarily true that a null effect is indicative of no affective change, while a relatively positive or negative judgement bias does not necessarily correspond to a relatively positive or negative affective state.

When interpreting the results of judgement bias studies it is important to consider the impact of affect manipulation on reward and punisher valuation, and the direction of this effect. Additionally, consideration should be given to within-test reward and punisher experience. It is evident from our studies, albeit with more drastic changes in within-test experience than standard judgement bias tests as a result of the fluctuating reward magnitude, that the ongoing rewards and punisher experience within the test session can influence decision-making. Importantly, affective valence might indirectly influence judgement bias by modulating the extent to which reward and punisher experience informs decision-making, leading to shifts in experience-dependent variability, as opposed to absolute shifts, in judgement bias.

Computational modelling provides a means for these potential issues to be addressed. In this thesis we provided a novel model for investigating judgement bias, which permitted insights into the relationship between reward and punisher experience, affect, and decision-making that could not be achieved through a statistical analysis alone. In particular, it allowed us to assess precisely which aspects of reward and punisher experience altered decision-making and affect and elucidated the mechanisms by which reward and punisher experience altered decision-making. Additional behavioural tests, such the progressive ratio task, also provided a method to directly assess the influence of the affect manipulation on reward valuation. A checklist of potential factors, other than an altered reward/punisher probability estimates, such as reward/punisher valuation or variation in within-test responses to rewards and punishers, that may influence judgement bias and which should be considered in the interpretation of results, or addressed using computational modelling or additional behaviour tests, may be useful to researchers conducting judgement bias tests.

In this thesis, we also describe a novel task which could provide an alternative measure of affective valence that requires a much shorter training period and use of a simpler computational model for data analysis. The task examines the point at which a rat opts to forgo a diminishing reward in order to receive a larger reward by providing two spatially distinct food sources which deplete following consecutive visits and restore following a visit to the alternate food source. Using this task, we found that rats in putatively negative affective states, resulting from following removal of

environmental enrichment from their homecage, would more readily forgo a reward, which was identified using both a statistical and computational analysis. In particular, this appeared to reflect greater stochasticity in decision-making, which could result from a reduced subjective value of the reward. Further research will be needed to validate this task and assess the extent to which it provides a measure of anxiety-like states as well as depression-like states.

8.5 Limitations and outstanding questions

One major limitation of this thesis is that comparing laboratory rats and humans, even when both complete near-identical tasks, is perhaps akin to comparing apples and oranges. As previously mentioned, laboratory rats are a very homogeneous population with standardised life experiences in controlled environments, who receive a substantial period of training, and (almost) no choice in whether they participate[2]. In contrast, humans have varied life experiences, a short period of training, and consent to participation. Moreover, in this thesis all rat experiments were conducted with male rats, whereas a majority (83.3%) of humans were female due to the unequal ratio of female to male staff and students at Bristol Veterinary School. Thus, while differences in results between rat and human studies could relate to species differences, it could also relate to other factors including environment, sex differences, training duration, or a self-selecting bias whereby personality traits such as extroversion or inquisitiveness determine whether a human decides to participant.

There is also an issue of generalisability; we cannot say whether the results from these studies generalise to all humans (given that all human participants lived in the UK with some connection to Bristol Vet School) or all rats (given that all rats were obtained from the same breeding facility). This is particularly relevant to the rat studies which had a very small sample size as a result of the rats failure to learn the task and equipment failure. Indeed, the results of the study in which the majority of rats failed to learn the task (Chapter 6) may simply reflect some characteristic of the rats associated with their ability to perform the task (e.g. rapid learners/good discrimination skills/highly food motivated). In future experiments, it would be interesting to run a judgement bias study with a more heterogeneous population of rats that include both males and females such as wild rats or even pet rats. Recently, Ellis et al. (2019) described a study in which pet rats participated in a behavioural experiment demonstrating that this would be feasible.

Additionally, although we attempted to design judgement bias tasks that were as similar as possible for both rats and humans, there were differences that should be resolved in future studies. The first, as previously mentioned, is that humans were shown the magnitude of the potential reward and loss prior to stimulus onset whereas rats were not made aware of the magnitude of the potential outcome. Although we

[2]Note that rats have been known to nap during training or testing and hence in these cases choose not to participate.

attempted to inform rats of the magnitude of the outcome by varying the amplitude of the tone and there was some evidence that rats did learn to discriminate between tone amplitude, performance at the low amplitude tones was very poor and we failed to train rats on this variant of the task.

The stimuli also differed between the studies; visual stimuli were used for human experiments and auditory stimuli were used for rat experiments. These stimuli were based both on constraints imposed by the testing equipment and also consideration of each species' perceptual systems. Rats are a nocturnal prey species, which is reflected in their visual system; approximately 99% of the photoreceptors in the rat retina are rods (responsible for vision in low light but colour insensitive; LaVail, 1976), compared with approximately 95% in humans (Curcio et al., 1990), and the position of their eyes on the side of their head allows a wide visual and overhead vision at the expense of depth perception (Wallace et al., 2013). However, rats have highly sensitive hearing and olfaction which is important for communication and predator detection (Blanchard et al., 2003; Doty, 1986; Knutson et al., 2002), and somatosensation (through their whiskers) which is essential for navigation, orientation, and balance (Arabzadeh et al., 2005; Mitchinson et al., 2007); the sensitivity of these senses in rats exceeds that of human senses (Guić-Robles et al., 1989; Kelly and Masterton, 1977; Rajan et al., 2006). The neural differences between humans and rats also make comparing decision-making between these species challenging. For example, the cerebral cortex of a rat is approximately a thousand times smaller than that of humans (Uylings and van Eden, 1991). Additionally, the distribution of monoamine transporters (e.g serotonin and norepinephrine) also differs between humans and rats (Smith and Porrino, 2008).

It is also important to highlight the short-term nature of the manipulations of reward and punisher experience described in this thesis, with the exception of the study described in Chapter 7 in which housing conditions were manipulated. The results relating to reward and punisher experience may reflect a transient affective response as opposed to longer-term mood. To better understand the long-term impact of reward and punisher experience on affect and decision-making it will be necessary to run a longitudinal study and monitor the impact of variable reward and punisher experience. In humans, this could be achieved through use of a smartphone-based platform through which participants regularly report on rewarding and punishing experiences and complete a short judgement bias task. Computational psychiatry research has already been conducted using a smartphone app for data collection with good success and large sample sizes (Rutledge et al., 2017). In rats, development of an automated judgement bias task which can be conducted within the homecage would be invaluable to permit regular judgement bias testing. Furthermore, research should be undertaken to investigate how decision-making on the judgement bias task might be altered in both human and a range of non-human animals experiencing affective disorders, for example asking whether long-term absolutely unfavourable environments might contribute to depression or anxiety, or asking how depression or anxiety might modulate

the relationship between unpredictability and decision-making on the judgement bias task. Indeed, the influence of prior beliefs on decision-making, and the relationship between these prior beliefs, affect, and reward and punisher experience, may become more apparent following studies which include participants experiencing affective disorders.

A key finding of this thesis is that variation in reward and punisher valuation partly underlies variation in judgement bias. A better understanding of the influence of affect on reward and punisher valuation, and particularly the direction of this effect, would aid the interpretation of judgement bias data and formulation of hypotheses for judgement bias testing. Hence, it would be worthwhile to examine how reward and punisher experience, both more and less favourable and both short and long term, influence reward and punisher valuation, using several measures of reward and punisher valuation in combination with computational modelling. It would also be interesting to examine the effect that reward and punisher controllability exerts on decision-making behaviour and how this interacts with predictability, given that controllability has been proposed to influence the affective impact of rewards and punishers (Bassett and Buchanan-Smith, 2007). For example, learnt helplessness, which is associated with a depression-like state, demonstrates that individuals will not try to escape a punisher in new contexts if they have previously learnt that punishers cannot be controlled (Seligman, 1972).

With regards to the computational model, it is important to bear in mind that a good model fit does not provide unequivocal evidence of the cognitive mechanisms underlying behaviour; alternate mechanisms that have not been modelled might instead account for behaviour. The putative decision-making processes characterised by our computational model were selected based on empirical and theoretical research; a wealth of research suggested that affect influences both reward and punisher probability estimation and valuation, and that affect should be influenced by reward and punisher experience, including predictability and absolute favourability (see Chapter 1). Although the computational model provided an overall good fit, model performance was noticeably poorer at the positive reference stimulus; humans participants exhibited greater accuracy than the model predicted. Thus, it is clear that our model does not fully account for human behaviour. Alongside the findings of the meta-analysis (Chapter 2) in which pharmacological manipulations were found to exert the weakest influence at the positive reference cue, this might suggest that responses to certain rewards (but not certain punishers) are impervious to the effects of previous experience and/or affect. In future, computational modelling could be used to further explore this result and to find a parameter that improves the model fit at the positive reference stimulus, with no detrimental effect at other stimulus levels. An important next step to assess our novel POMDP model will be to investigate how the current model parameters relate to neurobiological processes. Endogenous fluctuations in dopaminergic activity in the midbrain have been shown to correspond to within-subject variability in risky decision-making (Chew et al., 2019); hence we

might expect our results to correspond to the mean or standard deviation of these fluctuations within an individual participant.

As part of this thesis, only potential biases in decision-making have been investigated as there is clear evidence that affect can influence the different components of decision-making which can be easily investigated through behavioural tasks and computational modelling. However, an alternate or additional explanation is that biases in decision-making occur at the sensory level (Mendl et al., 2009). More specifically, the ambiguous stimuli could be perceived as either rewarding (i.e. 'seeing the world through rose-tinted glasses') or punishing, rather than perceived as ambiguous where the decision about the best course of action is determined by affect. There is some evidence to suggest that affect can alter perception. For example, anxiety is associated with a greater likelihood of perceiving an ambiguous facial expression as indicating a negative emotion (Blanchette et al., 2007). Although the extent to which affect may alter perception in a top-down manner has been debated and considered unlikely by some researchers (e.g. Firestone and Scholl, 2016), this warrants further investigation. Additionally, we have not examined the possibility that ambiguity aversion (Ellsberg, 1961) i.e. an aversion towards ambiguous outcomes and preference for known outcomes, might underlie our results. However, the meta-analysis revealed that judgement biases occur at non-ambiguous tones, suggesting that ambiguity aversion is unlikely to be the sole process underlying the relationship between affect and judgement bias.

8.6 Conclusions (see Fig. 8.1)

Through computational modelling, in particular use of a novel computational model, we have made progress in understanding the relationship between reward and punisher experience, affect, and decision-making. Although we have demonstrated, through a meta-analysis, that in general, putatively negative affect may result in a relatively negative judgement bias and vice versa, the relationship between affect and decision-making is complex. In particular, the aspects of reward and punisher experience which alter affect do not necessarily alter decision-making in humans, and affect may influence the extent to which decision-making is modulated by reward and punisher experience. Variation in judgement bias may not solely reflect changes in probability estimation but also reward or punisher valuation.

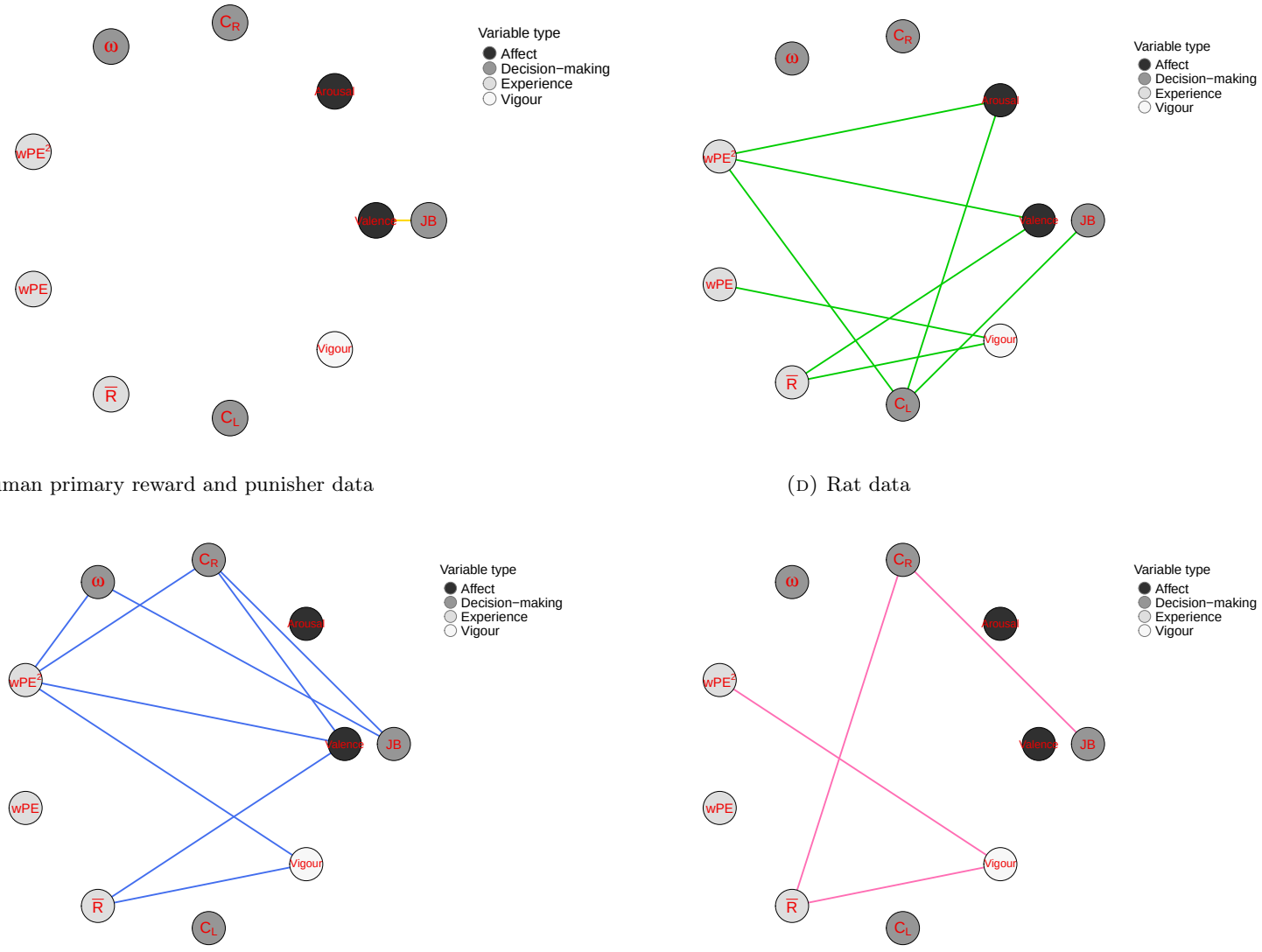
A key finding of this thesis is that reward and punisher experience influences (human) affect, (human and rat) vigour, and determines (human and rat) decision-making. More specifically, we have found that both rats and humans encode information about recent predictability and the absolute favourability of the environment to inform their behaviour. In humans, we also found that affect influenced decision-making by modulating the extent to which decision-making depended on the predictability of reward and punisher experience. Although the function of affect remains unclear, our results are most consistent with affect providing a measure of the

extent to which energy should be conserved to optimise the balance between reward acquisition, punisher avoidance, and energy expenditure given the environmental conditions.

In future studies, it would be worthwhile to further examine the hypotheses that a) negative affect is associated with less flexible model-free decision-making and b) reduced energy expenditure. This could be addressed by conducting experiments which allow model-free and model-based decision-making to be disentangled through computational modelling (e.g. Hasz and Redish, 2018, for rats; and Daw et al., 2011b, for humans) throughout which reward and/or punisher magnitude is systematically varied and human affect is measured as we have described in this thesis, and concurrent measures of energy expenditure (e.g. oxygen consumption, Jequier et al., 1987) are taken. Additionally, repeating the studies described in this thesis with a more heterogeneous population of rats and humans, as well as humans experiencing affective disorders, especially if conducted longitudinally, would provide further valuable insights into the adaptive function of affect across species. In particular, we might anticipate that predictability may be a more prominent factor in the decision-making process of rats who inhabit less predictable environments, and that the relationship between affect, environmental conditions, and probability estimation will be more apparent where there are greater differences in putative affective state between human (or rat) participants.

(A) Meta-analytic data

(B) Human secondary reward and loss data



(C) Human primary reward and punisher data

(D) Rat data

FIGURE 8.1: Diagrammatic summary of thesis results relating to the relationship between specific aspects of reward and punisher experience, affect, decision-making, and vigour: the nodes represent variables relating to affect, decision-making (where C_L denotes loss/punisher sensitivity, C_R denotes reward sensitivity, ω denotes the prior belief about outcomes, and JB denotes judgement bias), experience (where wPE denotes the weighted reward prediction error, $w\text{PE}^2$ denotes the squared weighted reward prediction error, and $\bar{R}$ denotes the average earning/reward rate), and vigour. The colour of the lines represents the study in which the association was observed.

Appendix A

Meta-analysis literature search

A.1 Main database search

A.1.1 Scopus

The following search was conducted using Scopus:

(TITLE-ABS-KEY (("Cognitive bias*" OR "judgement bias*" OR "judgement bias*" OR "Cognitive affective bias*") AND ("pessimis*" OR "optimis*" OR "valence" OR "mood*" OR "emotion*" OR "affective state*" OR "emotional state*" OR "ambig*") AND ("animal*" OR "animal welfare")) AND (PUBYEAR > 2016)) OR (TITLE-ABS-KEY (("Cognitive bias*" OR "judgement bias*" OR "judgement bias*" OR "Cognitive affective bias*") AND drug*))

A.1.2 Web of Science

The following search was conducted using Web of Science:

(TS=(("Cognitive bias*" OR "judgement bias*" OR "judgement bias*" OR "Cognitive affective bias*") AND ("pessimis*" OR "optimis*" OR "valence" OR "mood*" OR "emotion*" OR "affective state*" OR "emotional state*" OR "ambig*") AND ("animal*" OR "animal welfare")) AND (PY=(2017-2019))) OR (TS=(("Cognitive bias*" OR "judgement bias*" OR "judgement bias*" OR "Cognitive affective bias*") AND drug*)) Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC.

A.2 Subject database searches

The following databases were searched via the University of New South Wales:

A.2.1 Psychology databases

All APA Psychology databases via the Ovid platform including PsycINFO, PsycARTICLES, PsycBOOKS, PsycEXTRA and PsycTESTS using the following search:

(("Cognitive bias*" or "judgement bias*" or "judgement bias*" or "Cognitive affective bias*") and ("pessimis*" or "optimis*" or "valence" or "mood*" or "emotion*"

or "affective state*" or "emotional state*" or "ambig*") and ("animal*" or "animal welfare")).mp. [mp=ti, ab, td, hw, tc, id, ot, tm, mh, tx, ct]

A.2.2 Biomedical and pharmaceutical databases

The EMBASE database using the following search:

(("Cognitive bias*" or "judgement bias*" or "judgement bias*" or "Cognitive affective bias*") and ("pessimis*" or "optimis*" or "valence" or "mood*" or "emotion*" or "affective state*" or "emotional state*" or "ambig*") and ("animal*" or "animal welfare")).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]

The Medline database using the following search:

(("Cognitive bias*" or "judgement bias*" or "judgement bias*" or "Cognitive affective bias*") and ("pessimis*" or "optimis*" or "valence" or "mood*" or "emotion*" or "affective state*" or "emotional state*" or "ambig*") and ("animal*" or "animal welfare")).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]

A.3 Grey literature

The ProQuest Dissertation and Thesis database using the following searches:

noft(affect*) AND noft(bias*) AND noft(pharmacol*)

noft(affective) AND noft(bias*) AND noft(animal*) 20 records found

Google dataset using the following search:

("Cognitive bias*" OR "judgement bias*" OR "judgement bias*" OR "Cognitive affective bias*")

The Dimensions database using the following search:

affect* AND bias AND pharmacol* AND (animal OR welfare*)

A.4 Snowballing

The following review articles were used for snowballing using Scopus:

Crump, A., Arnott, G., and Bethell, E. (2018). Affect-driven attention biases as animal welfare indicators: review and methods. Animals, 8(8), 136.

Clegg, I. (2018). Cognitive bias in zoo animals: An optimistic outlook for welfare assessment. Animals, 8(7), 104.

Hales, C. A., Stuart, S. A., Anderson, M. H., and Robinson, E. S. (2014). Modelling cognitive affective biases in major depressive disorder using rodents. British journal of pharmacology, 171(20), 4524-4538.

Appendix B

POMDP Model Validation

To assess whether the model-fitting procedure could recover the true parameter values, we first generated sets of data using a range of parameter values for each parameter and fitted the model to these data. To determine whether the model could reliably detect individual differences, we examined the extent to which the true and recovered parameters were correlated using Pearson's correlation coefficient. Additionally, to determine the accuracy of parameter recovery, we examined the extent to which the mean of the true parameters differed from the mean of the recovered parameters using a t-test. Across all parameters (with the exception of $\gamma_{w\mathrm{PE}}$ for which there was a marginally non-significant correlation) there was a significant correlation (Fig. B.1, Table B.1) and no significant difference (Table B.1) between the true and recovered parameter values.

TABLE B.1: Results of the statistical analyses of the true and recovered parameter values for all parameters

Parameter	**Pearson's correlation coefficient**	**p-Value**	**Recovered parameter mean**	**True parameter mean**	**t-value**	**p-Value**
$\alpha_{\bar{R}}$	0.705	<0.001	-3.790	-4.007	-0.901	0.369
B	0.920	<0.001	0.697	0.681	-0.125	0.901
$\beta_0^{C_L}$	0.979	<0.001	0.440	0.462	0.135	0.892
$\beta_{\bar{R}}^{C_L}$	0.988	<0.001	0.113	0.095	-0.271	0.786
$\beta_O^{C_L}$	0.910	<0.001	0.020	0.021	0.279	0.780
$\beta_{w\mathrm{PE}}^{C_L}$	0.937	<0.001	0.066	0.042	-0.800	0.425
$\beta_{w\mathrm{PE}^2}^{C_L}$	0.927	<0.001	0.000	0.000	0.284	0.777
$\beta_0^{C_R}$	0.983	<0.001	0.483	0.471	-0.073	0.942
$\beta_{\bar{R}}^{C_R}$	0.943	<0.001	-0.100	-0.106	-0.177	0.859
$\beta_O^{C_R}$	0.874	<0.001	0.022	0.021	-0.159	0.874
$\beta_{w\mathrm{PE}}^{C_R}$	0.926	<0.001	0.016	0.016	0.035	0.972
$\beta_{w\mathrm{PE}^2}^{C_R}$	0.867	<0.001	0.000	0.000	0.318	0.751
β_0^{ω}	0.988	<0.001	-0.072	-0.088	-0.084	0.933
$\beta_{\bar{R}}^{\omega}$	0.989	<0.001	-0.044	-0.052	-0.114	0.909
β_O^{ω}	0.849	<0.001	0.023	0.022	-0.317	0.752
$\beta_{w\mathrm{PE}}^{\omega}$	0.735	<0.001	0.017	0.018	0.018	0.985
$\beta_{w\mathrm{PE}^2}^{\omega}$	0.846	<0.001	0.000	0.000	0.172	0.863
$\gamma_{w\mathrm{PE}}$	0.178	0.079	-0.645	-0.383	0.462	0.645
ζ	0.507	<0.001	2.462	2.335	-0.448	0.655
λ	0.739	<0.001	-4.242	-4.326	-0.481	0.631
σ	0.956	<0.001	-0.730	-0.724	0.070	0.944
ϕ	0.594	<0.001	-0.718	-0.540	0.981	0.328

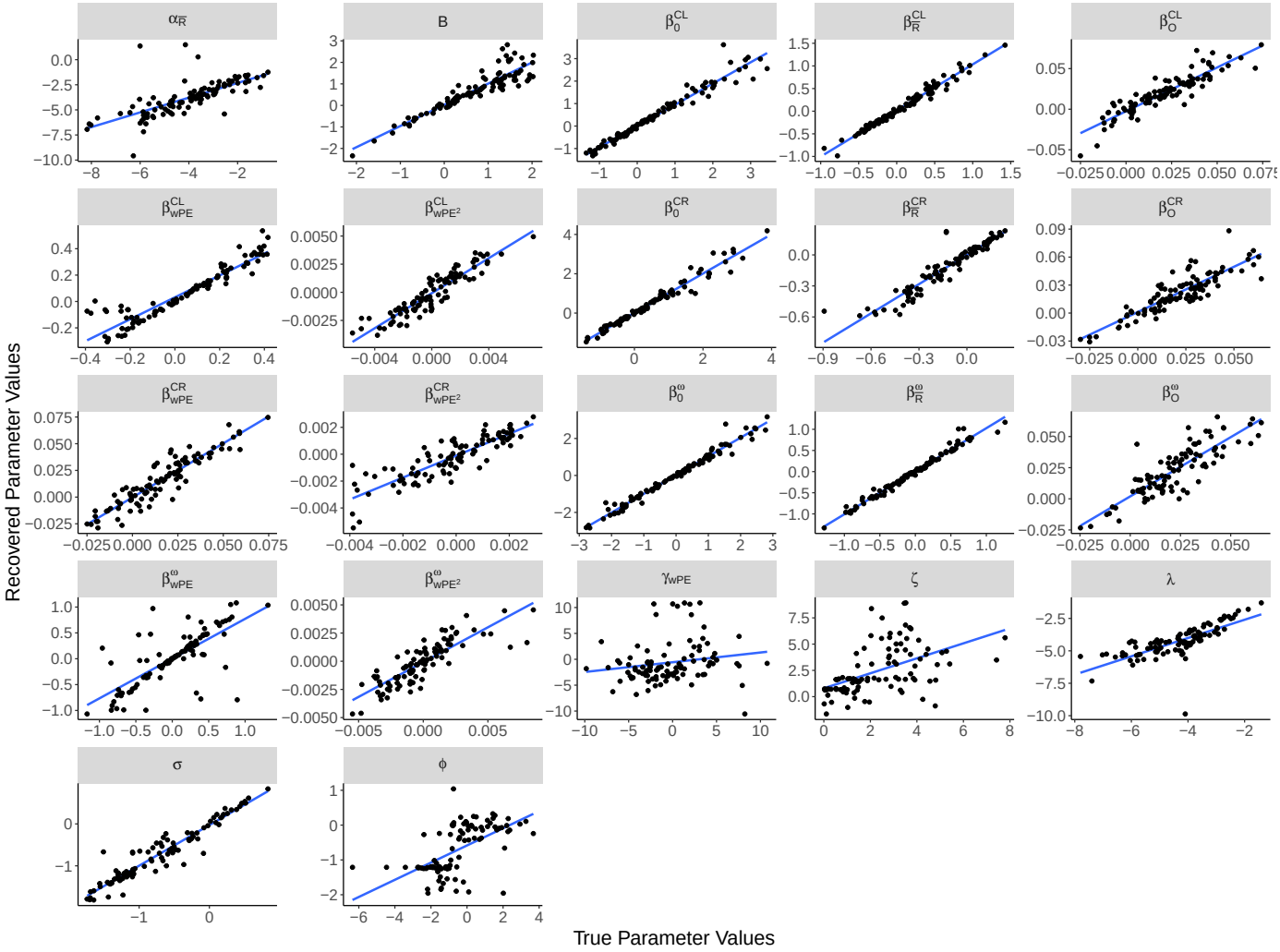


FIGURE B.1: Scatterplot of true and recovered parameter values for each parameter in POMDP model of judgement bias.

We then assessed the ability of the model-fitting procedure to recover the true parameter values from the best-fitting models from Chapters 3,4, and 6 when these parameters were fitted simultaneously. The parameter values used to generate data were within the range of parameters recovered from the observed data (i.e. drawn from a normal distribution with the mean and standard deviation of the recovered parameters). Each set of generated data had the same number of datapoints as the collected data (i.e. 180 for the studies with human subjects, and 90 for the rat study). We also assessed the extent to which there were correlations between the parameters recovered from the generated data, to observe the extent to which there were trade-offs between parameters.

There was a moderate correlation between the recovered parameter values of $\beta_0^{C_L}$ and σ, $\beta_0^{C_L}$ and λ, and $\beta_0^{C_L}$ and $\beta_{w\text{PE}^2}^{C_L}$ (Fig. B.2). Despite this, there was a significant correlation (Fig. B.3, Table B.2) and no significant difference (Table B.2) between the true and recovered parameter values. Hence, the parameters values obtained from the best-fitting model are considered to be reliable estimates of the true parameter

values.

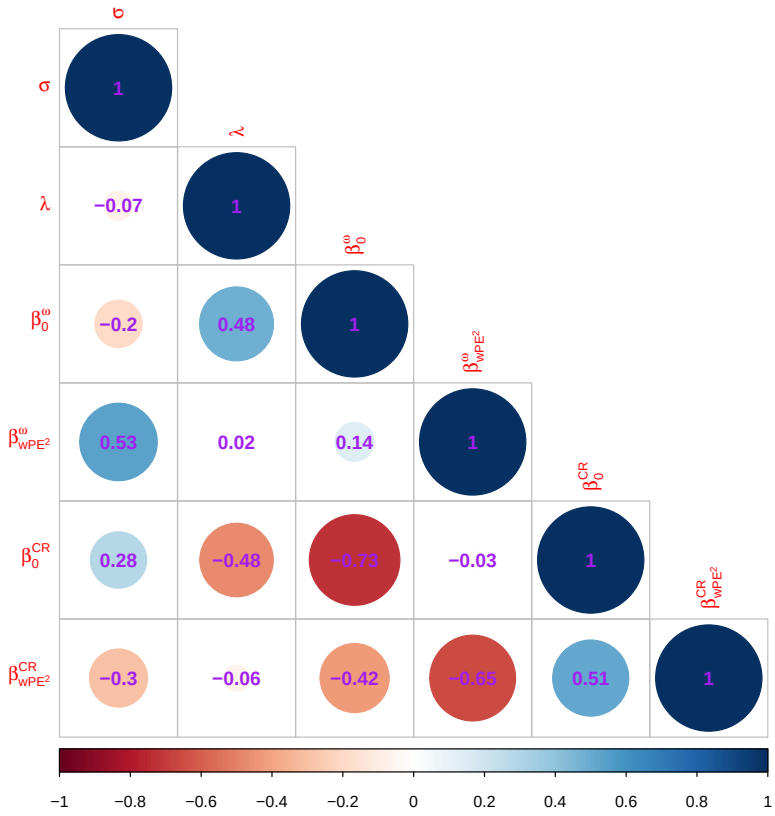


FIGURE B.2: Correlation matrix of recovered parameter values of σ, λ, $\beta_0^{C_L}$, $\beta_{wPE^2}^{C_L}$. The size of the circle indicates the magnitude of the correlation, the colour indicates both the magnitude and direction, and the number within the circle is the correlation coefficient.

TABLE B.2: Results of the statistical analyses of the true and recovered parameter values for parameters in the final model of data from the human monetary judgement bias task

Parameter	**Pearson's correlation coefficient**	**p-Value**	**Recovered parameter mean**	**True parameter mean**	**t-value**	**p-Value**
$\beta_0^{C_L}$	0.985	<0.001	-0.413	-0.448	-0.425	0.671
$\beta_{wPE_n^2}^{C_L}$	0.811	<0.001	0.002	0.003	0.022	0.982
λ	0.956	<0.001	-0.985	-1.007	-0.159	0.874
σ	0.524	<0.001	-0.559	-0.532	0.166	0.869

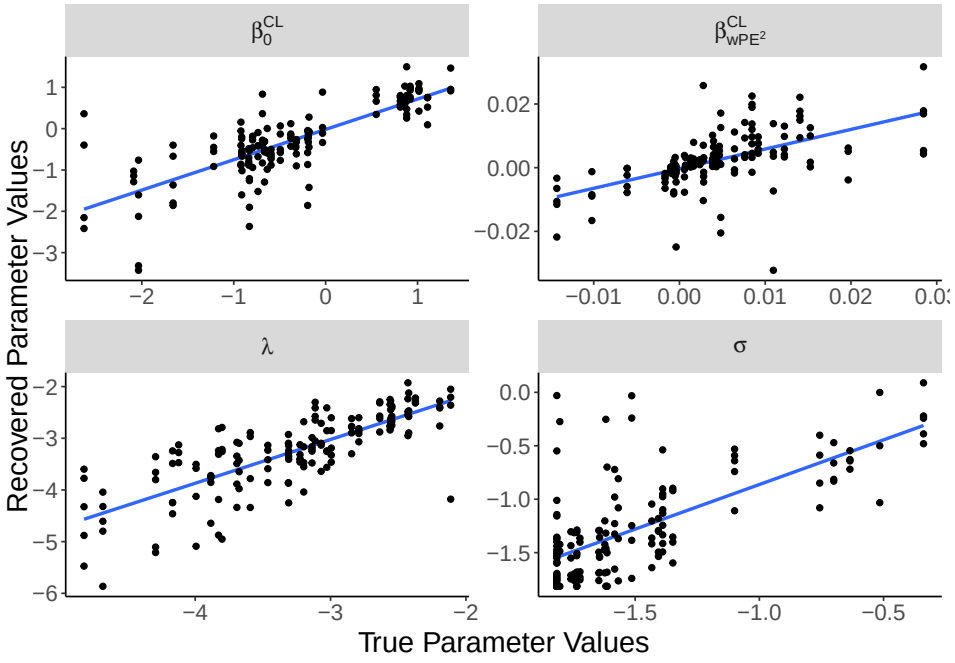


FIGURE B.3: Scatterplot of true and recovered parameter values for each parameter in the best-fitting model of data from the human monetary judgement bias task

There was a strong correlation between the recovered values of $\beta_0^{C_R}$ and β_0^{ω}, $\beta_{w\text{PE}^2}^{C_R}$ and $\beta_{w\text{PE}^2}^{\omega}$, $\beta_0^{C_R}$ and $\beta_{w\text{PE}^2}^{C_R}$, and σ and $\beta_{w\text{PE}^2}^{\omega}$ (Fig. B.4). Additionally, there was a moderate correlation between β_0^{ω} and λ, $\beta_0^{C_R}$ and λ, and β_0^{ω} and $\beta_{w\text{PE}^2}^{C_R}$ (Fig. B.4). Although there was a significant correlation between all true and recovered parameters (Fig. B.5, Table B.3), indicating that individual differences in these parameters were identifiable, there was a significant difference between the mean of the true and recovered values of β_0^{ω} (overestimate of true value), $\beta_0^{C_R}$ (underestimate of true value), and λ (underestimate of true value), see Table B.3. The estimates of these parameters and therefore not reliable. This could potentially be addressed in future studies using a greater number of trials or greater variation in the reward magnitude. However, this shortcoming does not alter the results qualitatively; if we consider that β_0^{ω} is an overestimate and $\beta_0^{C_R}$ our results would not change qualitatively, and no statistical tests were conducted on the value of λ.

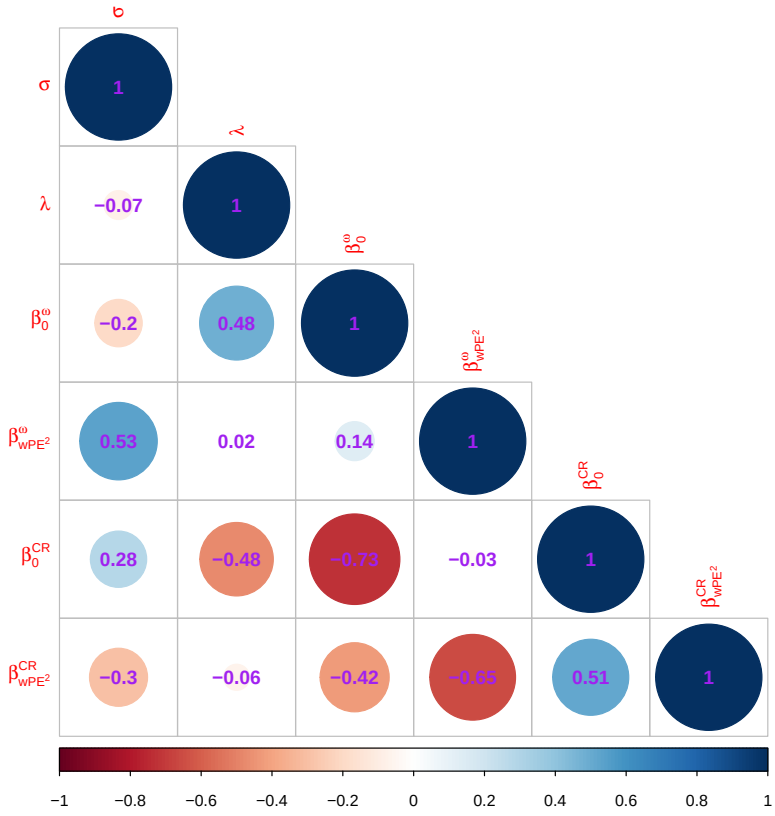


FIGURE B.4: Correlation matrix of recovered parameter values of σ, λ, β_0^{ω}, $\beta_{w\text{PE}^2}^{\omega}$, $\beta_0^{C_R}$, $\beta_{w\text{PE}^2}^{C_R}$. The size of the circle indicates the magnitude of the correlation, the colour indicates both the magnitude and direction, and the number within the circle is the correlation coefficient.

TABLE B.3: Results of the statistical analyses of the true and recovered parameter values for parameters in the final model of data from the human primary reward and punisher judgement bias task

Parameter	**Pearson's correlation coefficient**	**p-Value**	**Recovered parameter mean**	**True parameter mean**	**t-value**	**p-Value**
β_0^{ω}	0.413	<0.001	-1.476	-2.135	-3.039	0.003
$\beta_{w\text{PE}^2}^{\omega}$	0.321	<0.001	-0.051	-0.050	0.008	0.994
$\beta_0^{C_R}$	0.315	<0.001	1.513	2.294	3.804	0.000
$\beta_{w\text{PE}^2}^{C_R}$	0.484	<0.001	-0.162	-0.135	0.424	0.672
λ	0.327	<0.001	-12.591	-10.531	2.336	0.020
σ	0.771	<0.001	-1.552	-1.559	-0.161	0.872

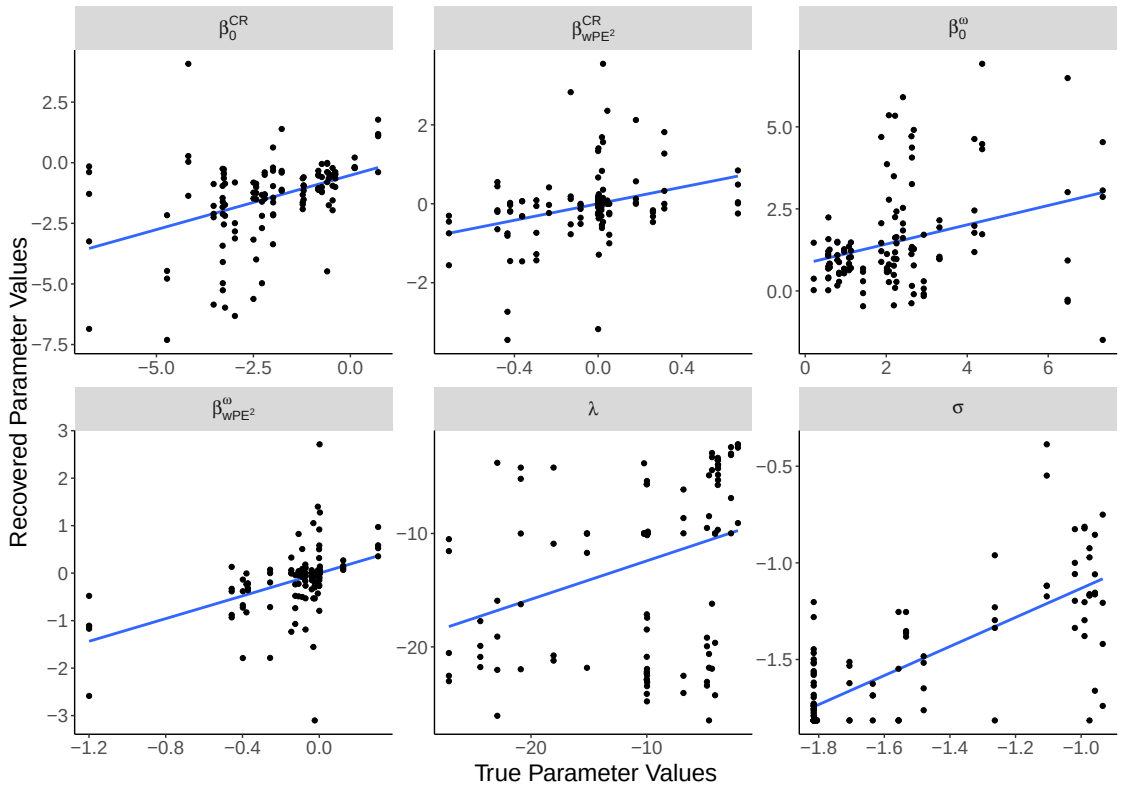


FIGURE B.5: Scatterplot of true and recovered parameter values for each parameter in the best-fitting model of data from the human primary reward and punisher judgement bias task

There was a strong correlation between $\beta_{\bar{R}}^{C_R}$ and $\alpha_{\bar{R}}$, and a moderate correlation between σ and all other parameters (i.e. $\beta_0^{C_R}$, $\beta_{\bar{R}}^{C_R}$, and $\alpha_{\bar{R}}$), see Fig. B.6. Recovery of $\alpha_{\bar{R}}$ was poor (Fig. B.7, Table B.4); the true and recovered parameters were not correlated, although there was no significant difference between the true and recovered means. Hence, inferences about variation across rats cannot be made using this parameter. The model fitting procedure underestimated the value of σ and tended to underestimate the value of $\beta_0^{C_R}$, although as above, this would not change the results qualitatively (Fig. B.7, Table B.4).

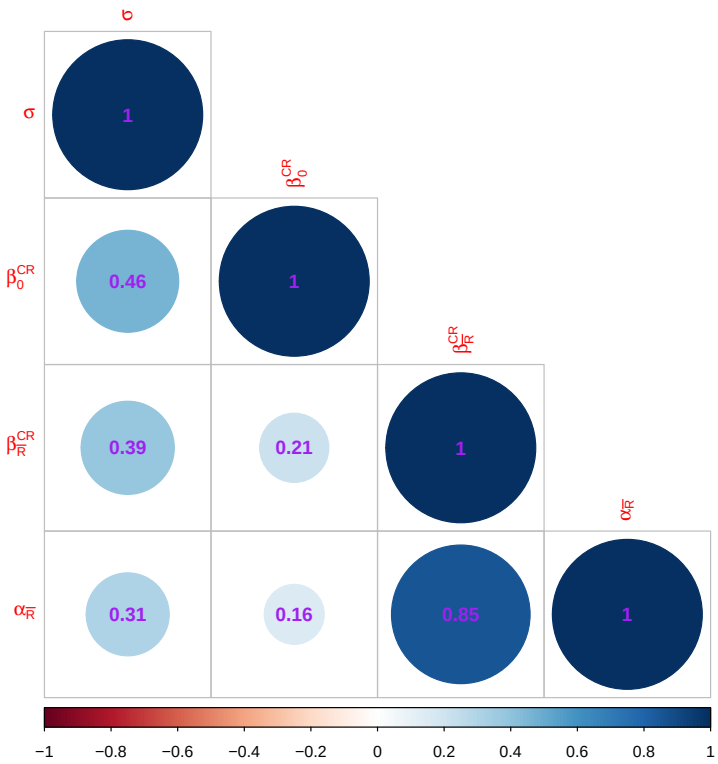


FIGURE B.6: Correlation matrix of recovered parameter values of $\alpha_{\bar{R}}$, $\beta_0^{C_R}$, $\beta_{\bar{R}}^{C_R}$, and σ. The size of the circle indicates the magnitude of the correlation, the colour indicates both the magnitude and direction, and the number within the circle is the correlation coefficient.

TABLE B.4: Results of the statistical analyses of the true and recovered parameter values for parameters of data from the rat judgement bias task

Parameter	**Pearson's correlation coefficient**	**p-Value**	**Recovered parameter mean**	**True parameter mean**	**t-value**	**p-Value**
$\alpha_{\bar{R}}$	-0.037	0.864	-2.723	-3.256	-1.215	0.231
$\beta_0^{C_R}$	0.720	$<$0.001	0.886	1.258	1.815	0.078
$\beta_{\bar{R}}^{C_R}$	0.455	0.025	0.368	0.182	-0.176	0.862
σ	0.604	0.002	-1.455	-0.986	3.176	0.003

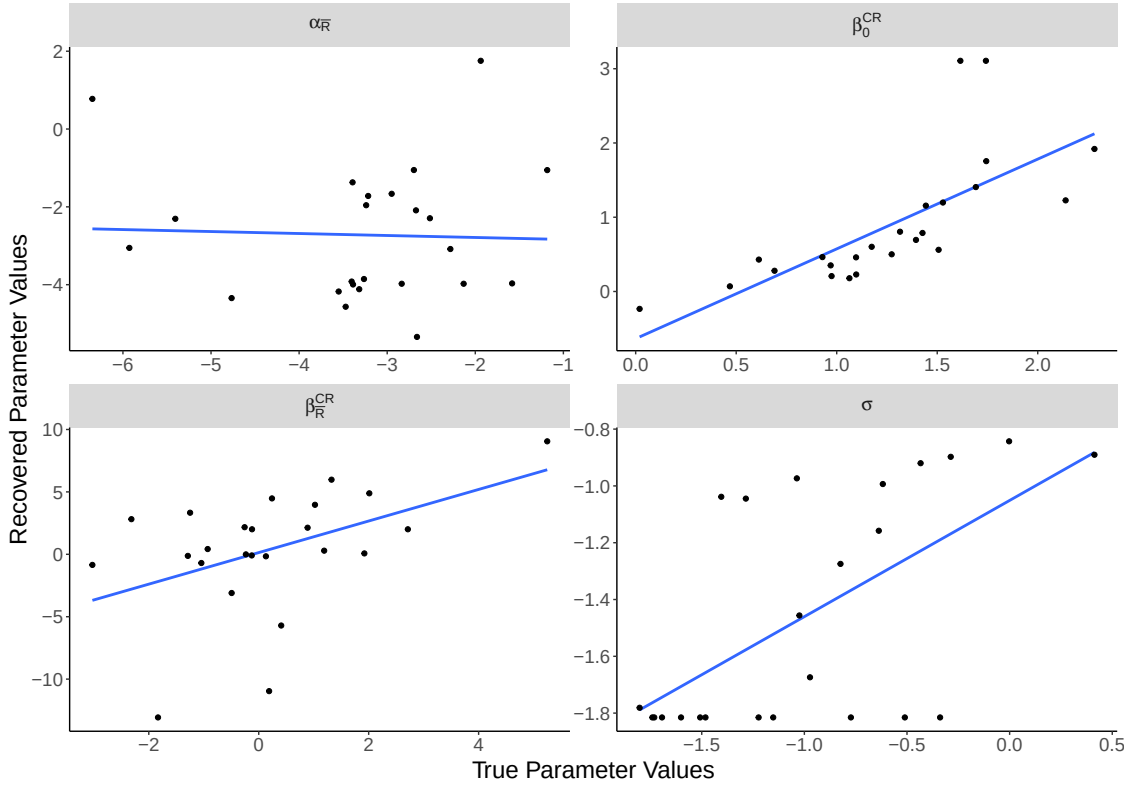


FIGURE B.7: Scatterplot of true and recovered parameter values for each parameter in the best-fitting model of data from the rat judgement bias task

It is important to note that data generated from values outside the range of values used here may not be recoverable. Indeed, the behavioural signature of extremely high or low values of several parameters (e.g. $\beta_0^{C_L}$, $\beta_0^{C_R}$, β_0^{ω}) would be near identical (i.e. a high proportion of 'stay' or 'go' responses across all ambiguous stimuli) and hence the model would unlikely be able to recover these values. Likewise, the model assumes generalisation of the stimuli (i.e. the summarised judgement bias data should follow an approximately sigmoidal shape when plotted), and consequently the model would struggle to capture behaviour in which generalisation was not observed. More specifically, although σ and λ govern the shape of the psychometric curve, there are no values at which these parameters could capture a non-monotonic curve (e.g. 100% accuracy at the most ambiguous cues; but 0% accuracy at the moderately ambiguous cues, and 100% accuracy and the non-ambiguous cues).

Bibliography

Abelson, J. L., Khan, S., Liberzon, I., and Young, E. A. (2007). HPA axis activity in patients with panic disorder: review and synthesis of four studies. *Depression and Anxiety*, 24(1):66–76.

Adams, R. A., Huys, Q. J., and Roiser, J. P. (2016). Computational psychiatry: towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery, and Psychiatry*, 87(1):53–63.

Addicott, M., Pearson, J., Sweitzer, M., Barack, D., and Platt, M. (2017). A primer on foraging and the explore/exploit trade-off for psychiatry research. *Neuropsychopharmacology*, 42(10):1931–1939.

Allritz, M., Call, J., and Borkenau, P. (2016). How chimpanzees (pan troglodytes) perform in a modified emotional stroop task. *Animal Cognition*, 19(3):435–449.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. American Psychiatric Association Arlington, VA, 5th edition.

Andersen, S. M., Spielman, L. A., and Bargh, J. A. (1992). Future-event schemas and certainty about the future: Automaticity in depressives' future-event predictions. *Journal of Personality and Social Psychology*, 63(5):711–723.

Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology: General*, 134(2):258–281.

Anderson, D. J. and Adolphs, R. (2014). A framework for studying emotions across species. *Cell*, 157(1):187–200.

Anderson, M. H., Hardcastle, C., Munafò, M. R., and Robinson, E. S. (2012). Evaluation of a novel translational task for assessing emotional biases in different species. *Cognitive, Affective, & Behavioral Neuroscience*, 12(2):373–381.

Arabzadeh, E., Zorzin, E., and Diamond, M. E. (2005). Neuronal encoding of texture in the whisker sensory pathway. *PLoS Biology*, 3(1):e17.

Arnone, D., Job, D., Selvaraj, S., Abe, O., Amico, F., Cheng, Y., Colloby, S. J., O'Brien, J. T., Frodl, T., Gotlib, I. H., et al. (2016). Computational meta-analysis of statistical parametric maps in major depression. *Human Brain Mapping*, 37(4):1393–1404.

Arnone, D., McKie, S., Elliott, R., Juhász, G., Thomas, E., Downey, D., Williams, S., Deakin, J., and Anderson, I. (2013). State-dependent changes in hippocampal grey matter in depression. *Molecular Psychiatry*, 18(12):1265–1272.

Åström, K. J. (1965). Optimal control of markov processes with incomplete state information. *Journal of Mathematical Analysis and Applications*, 10(1):174–205.

Bach, D. R. and Dayan, P. (2017). Algorithms for survival: a comparative perspective on emotions. *Nature Reviews Neuroscience*, 18(5):311–319.

Baciadonna, L. and McElligott, A. G. (2015). The use of judgement bias to assess welfare in farm livestock. *Animal Welfare*, 24(1):81–91.

Bardo, M., Klebaur, J., Valone, J., and Deaton, C. (2001). Environmental enrichment decreases intravenous self-administration of amphetamine in female and male rats. *Psychopharmacology*, 155(3):278–284.

Barnett, J. L. and Hemsworth, P. H. (1990). The validity of physiological and behavioural measures of animal welfare. *Applied Animal Behaviour Science*, 25(1-2):177–187.

Barrett, L. F. (2014). The conceptual act theory: A précis. *Emotion Review*, 6(4):292–297.

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, 12(1):1–23.

Barrett, L. F. and Russell, J. A. (1999). The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science*, 8(1):10–14.

Bassett, L. and Buchanan-Smith, H. M. (2007). Effects of predictability on the welfare of captive animals. *Applied Animal Behaviour Science*, 102(3-4):223–245.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, 76(4):695–711.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Bateson, M. (2016). Optimistic and pessimistic biases: a primer for behavioural ecologists. *Current Opinion in Behavioral Sciences*, 12:115–121.

Bateson, M., Brilot, B., and Nettle, D. (2011). Anxiety: an evolutionary approach. *The Canadian Journal of Psychiatry*, 56(12):707–715.

Bateson, M. and Poirier, C. (2019). Can biomarkers of biological age be used to assess cumulative lifetime experience. *Animal Welfare*, pages 41–56.

Bechara, A. and Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, 52(2):336–372.

Bechara, A., Damasio, H., Damasio, A. R., and Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience*, 19(13):5473–5481.

Beck, A. and Alford, B. (1967). *Depression: causes and treatment.* University of Pennsylvania Press, Pennsylvania.

Beck, A. T. (2002). Cognitive models of depression. *Clinical Advances in Cognitive Psychotherapy: Theory and Application*, 14(1):29–61.

Beck, A. T., Ward, C., Mendelson, M., Mock, J., and Erbaugh, J. (1961). Beck depression inventory (BDI). *Archives of General Psychiatry*, 4(6):561–571.

Beck, S. M., Locke, H. S., Savine, A. C., Jimura, K., and Braver, T. S. (2010). Primary and secondary rewards differentially modulate neural activity dynamics during working memory. *PloS One*, 5(2):e9251.

Beierholm, U., Guitart-Masip, M., Economides, M., Chowdhury, R., Düzel, E., Dolan, R., and Dayan, P. (2013). Dopamine modulates reward-related vigor. *Neuropsychopharmacology*, 38(8):1495–1503.

Bell, P. A. and Baron, R. A. (1977). Aggression and ambient temperature: The facilitating and inhibiting effects of hot and cold environments. *Bulletin of the Psychonomic Society*, 9(6):443–445.

Bellman, R. (1952). On the theory of dynamic programming. *Proceedings of the National Academy of Sciences of the United States of America*, 38(8):716–719.

Bellman, R. (1957). A markovian decision process. *Journal of Mathematics and Mechanics*, pages 679–684.

Berlin, I., Givry-Steiner, L., Lecrubier, Y., and Puech, A. (1998). Measures of anhedonia and hedonic responses to sucrose in depressive and schizophrenic patients in comparison with healthy subjects. *European Psychiatry*, 13(6):303–309.

Bernoulli, D. (1954). Originally published in 1738; translated by L. Sommer. exposition of a new theory on the measurement of risk. *Econometrica*, 22(1):22–36.

Berridge, K. C. (1996). Food reward: brain substrates of wanting and liking. *Neuroscience & Biobehavioral Reviews*, 20(1):1–25.

Berridge, K. C. (2012). From prediction error to incentive salience: mesolimbic computation of reward motivation. *European Journal of Neuroscience*, 35(7):1124–1143.

Berridge, K. C. and Aldridge, J. W. (2008). Special review: Decision utility, the brain, and pursuit of hedonic goals. *Social Cognition*, 26(5):621–646.

Berridge, K. C. and Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, 28(3):309–369.

Berridge, K. C. and Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences*, 26(9):507–513.

Berridge, K. C., Robinson, T. E., and Aldridge, J. W. (2009). Dissecting components of reward:'liking','wanting', and learning. *Current Opinion in Pharmacology*, 9(1):65–73.

Bestmann, S., Harrison, L. M., Blankenburg, F., Mars, R. B., Haggard, P., Friston, K. J., and Rothwell, J. C. (2008). Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. *Current Biology*, 18(10):775–780.

Bethell, E. J. (2015). A "how-to" guide for designing judgment bias studies to assess captive animal welfare. *Journal of Applied Animal Welfare Science*, 18(S1):S18–S42.

Bethell, E. J., Holmes, A., MacLarnon, A., and Semple, S. (2012). Evidence that emotion mediates social attention in rhesus macaques. *PLoS One*, 7(8):e44387.

Bethell, E. J. and Koyama, N. F. (2015). Happy hamsters? Enrichment induces positive judgement bias for mildly (but not truly) ambiguous cues to reward and punishment in mesocricetus auratus. *Royal Society Open Science*, 2(7):140399.

Bezard, E., Dovero, S., Belin, D., Duconger, S., Jackson-Lewis, V., Przedborski, S., Piazza, P. V., Gross, C. E., and Jaber, M. (2003). Enriched environment confers resistance to 1-methyl-4-phenyl-1, 2, 3, 6-tetrahydropyridine and cocaine: involvement of dopamine transporter and trophic factors. *Journal of Neuroscience*, 23(35):10999–11007.

Bjärehed, J., Sarkohi, A., and Andersson, G. (2010). Less positive or more negative? Future-directed thinking in mild to moderate depression. *Cognitive Behaviour Therapy*, 39(1):37–45.

Blanchard, D. C., Griebel, G., and Blanchard, R. J. (2003). Conditioning and residual emotionality effects of predator stimuli: some reflections on stress and emotion. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 27(8):1177–1185.

Blanchette, I., Richards, A., and Cross, A. (2007). Anxiety and the interpretation of ambiguous facial expressions: The influence of contextual cues. *The Quarterly Journal of Experimental Psychology*, 60(8):1101–1115.

Blanco, N. J., Otto, A. R., Maddox, W. T., Beevers, C. G., and Love, B. C. (2013). The influence of depression symptoms on exploratory decision-making. *Cognition*, 129(3):563–568.

Blaney, P. H. (1986). Affect and memory: a review. *Psychological Bulletin*, 99(2):229–246.

Bogdanova, O. V., Kanekar, S., D'Anci, K. E., and Renshaw, P. F. (2013). Factors influencing behavior in the forced swim test. *Physiology & Behavior*, 118:227–239.

Bonnelle, V., Veromann, K.-R., Heyes, S. B., Sterzo, E. L., Manohar, S., and Husain, M. (2015). Characterization of reward and effort mechanisms in apathy. *Journal of Physiology-Paris*, 109(1-3):16–26.

Bora, E., Harrison, B., Yücel, M., and Pantelis, C. (2013). Cognitive impairment in euthymic major depressive disorder: a meta-analysis. *Psychological Medicine*, 43(10):2017–2026.

Borsini, F. and Meli, A. (1988). Is the forced swimming test a suitable model for revealing antidepressant activity? *Psychopharmacology*, 94(2):147–160.

Borsini, F., Podhorna, J., and Marazziti, D. (2002). Do animal models of anxiety predict anxiolytic-like effects of antidepressants? *Psychopharmacology*, 163(2):121–141.

Boureau, Y.-L. and Dayan, P. (2011). Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology*, 36(1):74–97.

Bouton, M. E. and King, D. A. (1983). Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *Journal of Experimental Psychology: Animal Behavior Processes*, 9(3):248–265.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., and Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, 30(2):619–639.

Briefer, E. F. and McElligott, A. G. (2013). Rescued goats at a sanctuary display positive mood after former neglect. *Applied Animal Behaviour Science*, 146(1-4):45–55.

Brilot, B. O., Asher, L., and Bateson, M. (2010). Stereotyping starlings are more 'pessimistic'. *Animal Cognition*, 13(5):721–731.

Broom, D. M. (1991). Animal welfare: concepts and measurement. *Journal of Animal Science*, 69(10):4167–4175.

Broom, D. M. and Johnson, K. G. (1993). *Stress and animal welfare*. Kluwer Academic Publishers, The Netherlands.

Brown, P. L. and Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, 11(1):1–8.

Browning, M., Behrens, T. E., Jocham, G., O'reilly, J. X., and Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, 18(4):590.

Brydges, N. M., Leach, M., Nicol, K., Wright, R., and Bateson, M. (2011). Environmental enrichment induces optimistic cognitive bias in rats. *Animal Behaviour*, 81(1):169–175.

Buerkle, H. and Yaksh, T. (1998). Pharmacological evidence for different alpha 2-adrenergic receptor sites mediating analgesia and sedation in the rat. *British Journal of Anaesthesia*, 81(2):208–215.

Burgdorf, J., Kroes, R. A., Moskal, J. R., Pfaus, J. G., Brudzynski, S. M., and Panksepp, J. (2008). Ultrasonic vocalizations of rats (rattus norvegicus) during mating, play, and aggression: Behavioral concomitants, relationship to reward, and self-administration of playback. *Journal of Comparative Psychology*, 122(4):357–367.

Burman, O., McGowan, R., Mendl, M., Norling, Y., Paul, E., Rehn, T., and Keeling, L. (2011). Using judgement bias to measure positive affective state in dogs. *Applied Animal Behaviour Science*, 132(3-4):160–168.

Burman, O. H., Parker, R., Paul, E. S., and Mendl, M. (2008a). A spatial judgement task to determine background emotional state in laboratory rats, rattus norvegicus. *Animal Behaviour*, 76(3):801–809.

Burman, O. H. P., Parker, R. M., Paul, E. S., and Mendl, M. (2008b). Sensitivity to reward loss as an indicator of animal emotion and welfare. *Biology Letters*, 4:330–333.

Butler, G. and Mathews, A. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy*, 5(1):51–62.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., and Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5):365.

Bylsma, L. M., Morris, B. H., and Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clinical Psychology Review*, 28(4):676–691.

Cabanac, M. (1992). Pleasure: the common currency. *Journal of Theoretical Biology*, 155(2):173–200.

Cahill, L. and McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neurosciences*, 21(7):294–299.

Calabrese, E. J. and Baldwin, L. A. (2001a). The frequency of u-shaped dose responses in the toxicological literature. *Toxicological Sciences*, 62(2):330–338.

Calabrese, E. J. and Baldwin, L. A. (2001b). Hormesis: U-shaped dose responses and their centrality in toxicology. *Trends in Pharmacological Sciences*, 22(6):285–291.

Calabrese, E. J. and Baldwin, L. A. (2001c). U-shaped dose-responses in biology, toxicology, and public health. *Annual review of public health*, 22(1):15–33.

Carmel, D. and Markovitch, S. (1990). Exploration strategies for model-based learning in multi-agent systems. *Autonomous Agents and Multi-agent Systems*, 2:173–207.

Carobrez, A. and Bertoglio, L. (2005). Ethological and temporal analyses of anxiety-like behavior: the elevated plus-maze model 20 years on. *Neuroscience & Biobehavioral Reviews*, 29(8):1193–1205.

Carter, C. (2001). Is there a neurobiology of good welfare. *Coping with Challenge: Welfare of Animals including Humans*, pages 11–30.

Carver, C. S. (2001). Affect and the functional bases of behavior: On the dimensional structure of affective experience. *Personality and Social Psychology Review*, 5(4):345–356.

Cassini, M. H., Kacelnik, A., and Segura, E. T. (1990). The tale of the screaming hairy armadillo, the guinea pig and the marginal value theorem. *Animal Behaviour*, 39(6):1030–1050.

Center for Drug Evaluation and Research (2019). FDA Adverse Event Reporting System (FAERS) Public Dashboard.

Charnov, E. L. et al. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, 9:129–136.

Charpentier, C. J., Aylward, J., Roiser, J. P., and Robinson, O. J. (2017). Enhanced risk aversion, but not loss aversion, in unmedicated pathological anxiety. *Biological Psychiatry*, 81(12):1014–1022.

Chase, H., Frank, M., Michael, A., Bullmore, E., Sahakian, B., and Robbins, T. (2010). Approach and avoidance learning in patients with major depression and healthy controls: relation to anhedonia. *Psychological Medicine*, 40(3):433–440.

Chew, B., Hauser, T. U., Papoutsi, M., Magerkurth, J., Dolan, R. J., and Rutledge, R. B. (2019). Endogenous fluctuations in the dopaminergic midbrain drive behavioral choice variability. *Proceedings of the National Academy of Sciences*, page 201900872.

Christianson, S.-Å. (1992). Emotional stress and eyewitness memory: a critical review. *Psychological Bulletin*, 112(2):284–309.

Chung, D., Kadlec, K., Aimone, J. A., McCurry, K., King-Casas, B., and Chiu, P. H. (2017). Valuation in major depression is intact and stable in a non-learning environment. *Scientific Reports*, 7:44374.

Cimadevilla, J. M., Fenton, A. A., and Bures, J. (2001). New spatial cognition tests for mice: passive place avoidance on stable and active place avoidance on rotating arenas. *Brain Research Bulletin*, 54(5):559–563.

Clark, J. E., Watson, S., and Friston, K. J. (2018). What is mood? A computational perspective. *Psychological Medicine*, 48(14):2277–2284.

Clifford, C. W. (2002). Perceptual adaptation: motion parallels orientation. *Trends in Cognitive Sciences*, 6(3):136–143.

Clore, G. L., Schwarz, N., and Conway, M. (1994). Affective causes and consequences of social information processing. *Handbook of social cognition*, 1:323–417.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Erlbaum Associates, Hillsdale, 2 edition.

Constantino, S. M. and Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective, & Behavioral Neuroscience*, 15(4):837–853.

Cornwell, B. R., Garrido, M. I., Overstreet, C., Pine, D. S., and Grillon, C. (2017). The unpredictive brain under threat: A neurocomputational account of anxious hypervigilance. *Biological Psychiatry*, 82(6):447–454.

Crozier, W. J. and Pincus, G. (1928). On the geotropic orientation of young mammals. *The Journal of General Physiology*, 11(6):789–802.

Crump, A., Arnott, G., and Bethell, E. (2018). Affect-driven attention biases as animal welfare indicators: review and methods. *Animals*, 8(8):136.

Cryan, J. F., Markou, A., and Lucki, I. (2002). Assessing antidepressant activity in rodents: recent developments and future needs. *Trends in Pharmacological Sciences*, 23(5):238–245.

Cryan, J. F. and Mombereau, C. (2004). In search of a depressed mouse: utility of models for studying depression-related behavior in genetically modified mice. *Molecular Psychiatry*, 9(4):326–357.

Curcio, C. A., Sloan, K. R., Kalina, R. E., and Hendrickson, A. E. (1990). Human photoreceptor topography. *Journal of Comparative Neurology*, 292(4):497–523.

Cussen, V. A. and Mench, J. A. (2014). Personality predicts cognitive bias in captive psittacines, amazona amazonica. *Animal behaviour*, 89:123–130.

Damasio, A. R. (1994). *Descartes' error: Emotion, rationality and the human brain*. Random House, London.

Darrow, S. M., Verhoeven, J. E., Révész, D., Lindqvist, D., Penninx, B. W., Delucchi, K. L., Wolkowitz, O. M., and Mathews, C. A. (2016). The association between psychiatric disorders and telomere length: a meta-analysis involving 14,827 persons. *Psychosomatic Medicine*, 78(7):776–787.

Darwin, C. (1872). *The expression of the emotions in man and animals*. Appleton, New York.

Daston, L. J. (1980). Probabilistic expectation and rationality in classical probability theory. *Historia Mathematica*, 7(3):234–260.

Davis, M., Walker, D. L., Miles, L., and Grillon, C. (2010). Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology*, 35(1):105–135.

Daw, N. D. et al. (2011a). Trial-by-trial data analysis using computational models. *Decision making, affect, and learning: Attention and performance XXIII*, 23(1).

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., and Dolan, R. J. (2011b). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69(6):1204–1215.

Dawkins, M. S. (2008). The science of animal suffering. *Ethology*, 114(10):937–945.

Dayan, P. and Berridge, K. C. (2014). Model-based and model-free pavlovian reward learning: revaluation, revision, and revelation. *Cognitive, Affective, & Behavioral Neuroscience*, 14(2):473–492.

Dayan, P., Hinton, G. E., Neal, R. M., and Zemel, R. S. (1995). The helmholtz machine. *Neural computation*, 7(5):889–904.

Dayan, P. and Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Current Opinion in Neurobiology*, 18(2):185–196.

De Vry, J. and Schreiber, R. (1997). The chronic mild stress depression model: future developments from a drug discovery perspective. *Psychopharmacology*, 134(4):349–350.

De Wit, S., Standing, H. R., DeVito, E. E., Robinson, O. J., Ridderinkhof, K. R., Robbins, T. W., and Sahakian, B. J. (2012). Reliance on habits at the expense of goal-directed control following dopamine precursor depletion. *Psychopharmacology*, 219(2):621–631.

Deakin, A., Browne, W. J., Hodge, J. J., Paul, E. S., and Mendl, M. (2016). A screen-peck task for investigating cognitive bias in laying hens. *PloS One*, 11(7):e0158222.

Delgado, M. R., Jou, R. L., and Phelps, E. A. (2011). Neural systems underlying aversive conditioning in humans with primary and secondary reinforcers. *Frontiers in Neuroscience*, 5:71.

Descovich, K., Wathan, J., Leach, M. C., Buchanan-Smith, H. M., Flecknell, P., Farningham, D., and Vick, S.-J. (2017). Facial expression: An under-utilised tool for the assessment of welfare in mammals. *ALTEX - Alternatives to Animal Experimentation*, 34(3):409–429.

Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., Costa, R. M., and Sousa, N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science*, 325(5940):621–625.

Dichter, G. S., Smoski, M. J., Kampov-Polevoy, A. B., Gallop, R., and Garbutt, J. C. (2010). Unipolar depression does not moderate responses to the sweet taste test. *Depression and Anxiety*, 27(9):859–863.

Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science*, 298(5596):1191–1194.

Doll, B. B., Simon, D. A., and Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology*, 22(6):1075–1081.

Dombrovski, A. Y., Clark, L., Siegle, G. J., Butters, M. A., Ichikawa, N., Sahakian, B. J., and Szanto, K. (2010). Reward/punishment reversal learning in older suicide attempters. *American Journal of Psychiatry*, 167(6):699–707.

Dombrovski, A. Y., Szanto, K., Clark, L., Reynolds, C. F., and Siegle, G. J. (2013). Reward signals, attempted suicide, and impulsivity in late-life depression. *JAMA psychiatry*, 70(10):1020–1030.

Doty, R. L. (1986). Odor-guided behavior in mammals. *Experientia*, 42(3):257–271.

Doyle, R. E., Fisher, A. D., Hinch, G. N., Boissy, A., and Lee, C. (2010). Release from restraint generates a positive judgement bias in sheep. *Applied Animal Behaviour Science*, 122(1):28–34.

Duncan, I. J. (2004). A concept of welfare based on feelings. In *The well-being of farm animals: Challenges and solutions*, pages 85–101. Blackwell, Iowa.

Dunn, B. D., Dalgleish, T., and Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience & Biobehavioral Reviews*, 30(2):239–271.

Dunning, D., Fetchenhauer, D., and Schlösser, T. (2017). The varying roles played by emotion in economic decision making. *Current Opinion in Behavioral Sciences*, 15:33–38.

Eastwood, J. D., Smilek, D., and Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception & Psychophysics*, 63(6):1004–1013.

Egger, M., Smith, G. D., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109):629–634.

Ekman, P. (1984). Expression and the nature of emotion. *Approaches to emotion*, 3:19–344.

Eldar, E., Rutledge, R. B., Dolan, R. J., and Niv, Y. (2016). Mood as representation of momentum. *Trends in Cognitive Sciences*, 20(1):15–24.

Elliott, R., Rubinsztein, J. S., Sahakian, B. J., and Dolan, R. J. (2002). The neural basis of mood-congruent processing biases in depression. *Archives of General Psychiatry*, 59(7):597–604.

Ellis, S. L., Riemer, S., Thompson, H., and Burman, O. H. (2019). Assessing the external validity of successive negative contrast–implications for animal welfare. *Journal of Applied Animal Welfare Science*, pages 1–8.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, pages 643–669.

Engelmann, M., Thrivikraman, K., Su, Y., Nemeroff, C., Montkowski, A., Landgraf, R., Holsboer, F., and Plotsky, P. (1996). Endocrine and behavioral effects of airpuff-startle in rats. *Psychoneuroendocrinology*, 21(4):391–400.

Eysenck, M. W., Mogg, K., May, J., Richards, A., and Mathews, A. (1991). Bias in interpretation of ambiguous sentences related to threat in anxiety. *Journal of Abnormal Psychology*, 100(2):144–150.

Finlayson, K., Lampe, J. F., Hintze, S., Würbel, H., and Melotti, L. (2016). Facial indicators of positive emotions in rats. *PloS One*, 11(11):e0166446.

Finucane, M. L., Alhakami, A., Slovic, P., and Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, 13(1):1–17.

Fiorillo, C. D., Tobler, P. N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, 299(5614):1898–1902.

Firestone, C. and Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, 39:e229.

Flaherty, C. F. (1982). Incentive contrast: A review of behavioral changes following shifts in reward. *Animal Learning & Behavior*, 10(4):409–440.

Foisy, M.-L., Kornreich, C., Fobe, A., D'hondt, L., Pelc, I., Hanak, C., Verbanck, P., and Philippot, P. (2007). Impaired emotional facial expression recognition in alcohol dependence: do these deficits persist with midterm abstinence? *Alcoholism: Clinical and Experimental Research*, 31(3):404–410.

Fontaine, J. R., Scherer, K. R., Roesch, E. B., and Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, 18(12):1050–1057.

Friston, K. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, 4(11):e1000211.

Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris*, 100(1-3):70–87.

Furukawa, T. A., McGuire, H., and Barbui, C. (2002). Meta-analysis of effects and side effects of low dosage tricyclic antidepressants in depression: systematic review. *BMJ*, 325:991–995.

Gibbon, J. (1977). Scalar expectancy theory and weber's law in animal timing. *Psychological Review*, 84(3):279.

Gibson, J. J. and Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. i. quantitative studies. *Journal of Experimental Psychology*, 20(5):453–467.

Gläscher, J., Daw, N., Dayan, P., and O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66(4):585–595.

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, 108(3):15647–15654.

Glimcher, P. W., Camerer, C. F., Fehr, E., and Poldrack, R. A. (2009). *Neuroeconomics: Decision Making and the Brain*. Academic Press.

Goddard, A. W., Mason, G. F., Almai, A., Rothman, D. L., Behar, K. L., Petroff, O. A., Charney, D. S., and Krystal, J. H. (2001). Reductions in occipital cortex GABA levels in panic disorder detected with 1h-magnetic resonance spectroscopy. *Archives of General Psychiatry*, 58(6):556–561.

Goldstein, D. S. and Kopin, I. J. (2008). Adrenomedullary, adrenocortical, and sympathoneural responses to stressors: a meta-analysis. *Endocrine Regulations*, 42(4):111.

Gradin, V. B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., and Steele, J. D. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain*, 134(6):1751–1764.

Green, D. M., Swets, J. A., et al. (1966). *Signal detection theory and psychophysics*, volume 1. Wiley, New York.

Green, T., Gehrke, B., and Bardo, M. (2002). Environmental enrichment decreases intravenous amphetamine self-administration in rats: dose-response functions for fixed-and progressive-ratio schedules. *Psychopharmacology*, 162(4):373–378.

Green, T. A., Cain, M. E., Thompson, M., and Bardo, M. T. (2003). Environmental enrichment decreases nicotine-induced hyperactivity in rats. *Psychopharmacology*, 170(3):235–241.

Griffiths, B. and Beierholm, U. R. (2017). Opposing effects of reward and punishment on human vigor. *Scientific Reports*, 7:42287.

Grillon, C. (2008). Models and mechanisms of anxiety: evidence from startle studies. *Psychopharmacology*, 199(3):421–437.

Grimm, J. W. and See, R. E. (2000). Dissociation of primary and secondary reward-relevant limbic nuclei in an animal model of relapse. *Neuropsychopharmacology*, 22(5):473.

Guić-Robles, E., Valdivieso, C., and Guajardo, G. (1989). Rats can learn a roughness discrimination using only their vibrissal system. *Behavioural Brain Research*, 31(3):285–289.

Guitart-Masip, M., Beierholm, U. R., Dolan, R., Duzel, E., and Dayan, P. (2011). Vigor in the face of fluctuating rates of reward: an experimental examination. *Journal of Cognitive Neuroscience*, 23(12):3933–3938.

Guitart-Masip, M., Huys, Q. J., Fuentemilla, L., Dayan, P., Duzel, E., and Dolan, R. J. (2012). Go and no-go learning in reward and punishment: interactions between affect and effect. *Neuroimage*, 62(1):154–166.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, 33(2):1–22.

Haenisch, B., Bilkei-Gorzo, A., Caron, M. G., and Bönisch, H. (2009). Knockout of the norepinephrine transporter and pharmacologically diverse antidepressants prevent behavioral and brain neurotrophin alterations in two chronic stress models of depression. *Journal of Neurochemistry*, 111(2):403–416.

Hale III, W. W., Jansen, J. H., Bouhuys, A. L., and Van den Hoofdakker, R. H. (1998). The judgment of facial expressions by depressed patients, their partners and controls. *Journal of Affective Disorders*, 47(1-3):63–70.

Hales, C. A., Robinson, E. S., and Houghton, C. J. (2016). Diffusion modelling reveals the decision making processes underlying negative judgement bias in rats. *PloS One*, 11(3):e0152592.

Hall, G. and Honey, R. C. (1989). Contextual effects in conditioning, latent inhibition, and habituation: associative and retrieval functions of contextual cues. *Journal of Experimental Psychology: Animal Behavior Processes*, 15(3):232–241.

Hamann, S. (2012). Mapping discrete and dimensional emotions onto the brain: controversies and consensus. *Trends in Cognitive Sciences*, 16(9):458–466.

Harding, E. J., Paul, E. S., and Mendl, M. (2004). Animal behaviour: cognitive bias and affective state. *Nature*, 427(6972):312.

Harlé, K. M., Guo, D., Zhang, S., Paulus, M. P., and Angela, J. Y. (2017). Anhedonia and anxiety underlying depressive symptomatology have distinct effects on reward-based decision-making. *PloS One*, 12(10):e0186473.

Hasz, B. M. and Redish, A. D. (2018). Deliberation and procedural automation on a two-step task for rats. *Frontiers in Integrative Neuroscience*, 12:39.

Hedges, L. V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *journal of Educational Statistics*, 6(2):107–128.

Higgins, J. P. and Green, S. (2011). Cochrane handbook for systematic reviews of interventions. *The Cochrane Collaboration*, pages 33–49.

Higgins, J. P. and Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11):1539–1558.

Hintze, S., Melotti, L., Colosio, S., Bailoo, J. D., Boada-Saña, M., Würbel, H., and Murphy, E. (2018). A cross-species judgement bias task: integrating active trial initiation into a spatial go/no-go task. *Scientific Reports*, 8(1):5104.

Hodos, W. (1961). Progressive Ratio as a Measure of Reward Strength. *Science*, 134(2):943–944.

Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal*, 50:346–363.

Hsee, C. K. and Rottenstreich, Y. (2004). Music, pandas, and muggers: on the affective psychology of value. *Journal of Experimental Psychology: General*, 133(1):23–30.

Hsee, C. K., Rottenstreich, Y., and Xiao, Z. (2005). When is more better? On the relationship between magnitude and subjective value. *Current Directions in Psychological Science*, 14(5):234–237.

Hughes, B. and Duncan, I. (1988). The notion of ethological 'need', models of motivation and animal welfare. *Animal Behaviour*, 36(6):1696–1707.

Hurnik, J. and Lehman, H. (1988). Ethics and farm animal welfare. *Journal of Agricultural Ethics*, 1(4):305–318.

Huys, Q. J. and Dayan, P. (2009). A bayesian formulation of behavioral control. *Cognition*, 113(3):314–328.

Huys, Q. J., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., and Roiser, J. P. (2012). Bonsai trees in your head: how the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Computational Biology*, 8(3):e1002410.

Huys, Q. J., Guitart-Masip, M., Dolan, R. J., and Dayan, P. (2015). Decision-theoretic psychiatry. *Clinical Psychological Science*, 3(3):400–421.

Huys, Q. J., Maia, T. V., and Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19(3):404–413.

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., and Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of Mood & Anxiety Disorders*, 3(1):12.

Huys, Q. J., Vogelstein, J., and Dayan, P. (2009). Psychiatry: Insights into depression through normative decision-making models. In *Advances in Neural Information Processing Systems*, pages 729–736.

Iigaya, K., Jolivald, A., Jitkrittum, W., Gilchrist, I. D., Dayan, P., Paul, E., and Mendl, M. (2016). Cognitive bias in ambiguity judgements: using computational models to dissect the effects of mild mood manipulation in humans. *PloS One*, 11(11):e0165840.

Ikemoto, S. and Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Research Reviews*, 31(1):6–41.

Inoue, T., Nakagawa, S., Izumi, T., Kitaichi, Y., and Koyama, T. (2006). Effect of combined treatment with noradrenaline and serotonin reuptake inhibitors on conditioned freezing. *European Journal of Pharmacology*, 540(1-3):91–95.

Isen, A. M., Nygren, T. E., and Ashby, F. G. (1988). Influence of positive affect on the subjective utility of gains and losses: it is just not worth the risk. *Journal of Personality and Social Psychology*, 55(5):710–717.

Isen, A. M. and Patrick, R. (1983). The effect of positive feelings on risk taking: When the chips are down. *Organizational Behavior and Human Performance*, 31(2):194–202.

James, W. (1884). What is an emotion? *Mind*, 9(34):188–205.

Jequier, E., Acheson, K., and Schutz, Y. (1987). Assessment of energy expenditure and fuel utilization in man. *Annual Review of Nutrition*, 7(1):187–208.

Joëls, M. and Baram, T. Z. (2009). The neuro-symphony of stress. *Nature Reviews Neuroscience*, 10(6):459–466.

Joffe, R. T., Bagby, R. M., Levitt, A. J., Regan, J. J., and Parker, J. D. (1993). The tridimensional personality questionnaire in major depression. *The American Journal of Psychiatry*, 150(6):959–960.

Johnson, E. J. and Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, 45(1):20–31.

Jones, S., Neville, V., Higgs, L., Paul, E. S., Dayan, P., Robinson, E. S., and Mendl, M. (2018). Assessing animal affect: an automated and self-initiated judgement bias task based on natural investigative behaviour. *Scientific Reports*, 8(1):12400.

Jorgensen, E. M. (2014). Animal evolution: looking for the first nervous system. *Current Biology*, 24(14):R655–R658.

Kacelnik, A. and Bateson, M. (1996). Risky theories—the effects of variance on foraging decisions. *American Zoologist*, 36(4):402–434.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Kalia, M. (2005). Neurobiological basis of depression: an update. *Metabolism*, 54(5):24–27.

Kambeitz, J. P. and Howes, O. D. (2015). The serotonin transporter in depression: Meta-analysis of in vivo and post mortem findings and implications for understanding and treating depression. *Journal of Affective Disorders*, 186:358–366.

Kara, N., Stukalin, Y., and Einat, H. (2017). Revisiting the validity of the mouse forced swim test: systematic review and meta-analysis of the effects of prototypic antidepressants. *Neuroscience & Biobehavioral Reviews*, 84:1–11.

Kato, T., Kubota, M., and Kasahara, T. (2007). Animal models of bipolar disorder. *Neuroscience & Biobehavioral Reviews*, 31(6):832–842.

Kelly, J. B. and Masterton, B. (1977). Auditory sensitivity of the albino rat. *Journal of comparative and physiological psychology*, 91(4):930.

Kermer, D. A., Driver-Linn, E., Wilson, T. D., and Gilbert, D. T. (2006). Loss aversion is an affective forecasting error. *Psychological Science*, 17(8):649–653.

Killgore, W. D. S. (1998). The affect grid: a moderately valid, nonspecific measure of pleasure and arousal. *Psychological reports*, 83(2):639–642.

Kleen, J. K., Sitomer, M. T., Killeen, P. R., and Conrad, C. D. (2006). Chronic stress impairs spatial memory and motivation for reward without disrupting motor ability and motivation to explore. *Behavioral Neuroscience*, 120(4):842.

Klimek, V., Stockmeier, C., Overholser, J., Meltzer, H. Y., Kalka, S., Dilley, G., and Ordway, G. A. (1997). Reduced levels of norepinephrine transporters in the locus coeruleus in major depression. *Journal of Neuroscience*, 17(21):8451–8458.

Knapp, G. and Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, 22(17):2693–2710.

Knill, D. C. and Pouget, A. (2004). The bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12):712–719.

Knutson, B., Burgdorf, J., and Panksepp, J. (2002). Ultrasonic vocalizations as indices of affective states in rats. *Psychological Bulletin*, 128(6):961–977.

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., and Glover, G. (2005). Distributed neural representation of expected value. *Journal of Neuroscience*, 25(19):4806–4812.

Koot, S., Baars, A., Hesseling, P., Van den Bos, R., and Joëls, M. (2013). Time-dependent effects of corticosterone on reward-based decision-making in a rodent model of the Iowa gambling task. *Neuropharmacology*, 70:306–315.

Krakenberg, V., Woigk, I., Rodriguez, L. G., Kästner, N., Kaiser, S., Sachser, N., and Richter, S. H. (2019). Technology or ecology? New tools to assess cognitive judgement bias in mice. *Behavioural Brain Research*, 362:279–287.

Krueger, T. H., Schedlowski, M., and Meyer, G. (2005). Cortisol and heart rate measures during casino gambling in relation to impulsivity. *Neuropsychobiology*, 52(4):206–211.

Kukolja, J., Schläpfer, T. E., Keysers, C., Klingmüller, D., Maier, W., Fink, G. R., and Hurlemann, R. (2008). Modeling a negative response bias in the human amygdala by noradrenergic–glucocorticoid interactions. *Journal of Neuroscience*, 28(48):12868–12876.

Kumar, P., Waiter, G., Ahearn, T., Milders, M., Reid, I., and Steele, J. (2008). Abnormal temporal difference reward-learning signals in major depression. *Brain*, 131(8):2084–2093.

Lane, R. D., Ahern, G. L., Schwartz, G. E., and Kaszniak, A. W. (1997). Is alexithymia the emotional equivalent of blindsight? *Biological Psychiatry*.

Lange, C. G. (1885). The mechanism of the emotions. *The Classical Psychologists*, pages 672–684.

Lau, B. and Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron*, 58(3):451–463.

LaVail, M. M. (1976). Rod outer segment disk shedding in rat retina: relationship to cyclic lighting. *Science*, 194(4269):1071–1074.

Laviola, G., Hannan, A. J., Macrì, S., Solinas, M., and Jaber, M. (2008). Effects of enriched environment on animal models of neurodegenerative diseases and psychiatric disorders. *Neurobiology of Disease*, 31(2):159–168.

Lazarou, J., Pomeranz, B. H., and Corey, P. N. (1998). Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA*, 279(15):1200–1205.

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press, New York.

Lee, C., Cafe, L. M., Robinson, S. L., Doyle, R. E., Lea, J. M., Small, A. H., and Colditz, I. G. (2018). Anxiety influences attention bias but not flight speed and crush score in beef cattle. *Applied Animal Behaviour Science*, 205:210–215.

Lee, C., Verbeek, E., Doyle, R., and Bateson, M. (2016). Attention bias to threat indicates anxiety differences in sheep. *Biology Letters*, 12(6):20150977.

Lee, R. S., Hermens, D. F., Porter, M. A., and Redoblado-Hodge, M. A. (2012). A meta-analysis of cognitive deficits in first-episode major depressive disorder. *Journal of Affective Disorders*, 140(2):113–124.

Lenow, J. K., Constantino, S. M., Daw, N. D., and Phelps, E. A. (2017). Chronic and acute stress promote overexploitation in serial decision making. *Journal of Neuroscience*, 37(23):5681–5689.

Leppänen, J. M., Milders, M., Bell, J. S., Terriere, E., and Hietanen, J. K. (2004). Depression biases the recognition of emotionally neutral faces. *Psychiatry Research*, 128(2):123–133.

Lin, Z. and Madras, B. (2006). Human genetics and pharmacology of neurotransmitter transporters. *Handbook of Experimental Pharmacology*, 175:327–371.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., and Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *Behavioral and Brain Sciences*, 35(3):121–143.

Loewenstein, G., Rick, S., and Cohen, J. D. (2008). Neuroeconomics. *Annual Review of Psychology*, 59(1):647–672.

Lövheim, H. (2012). A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical Hypotheses*, 78(2):341–348.

Ma, N. and Yu, A. J. (2015). Statistical learning and adaptive decision-making underlie human response time variability in inhibitory control. *Frontiers in Psychology*, 6:1046.

MacKay, D. J. and Mac Kay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.

MacLeod, A. K. and Byrne, A. (1996). Anxiety, depression, and the anticipation of future positive and negative experiences. *Journal of Abnormal Psychology*, 105(2):286–289.

MacLeod, A. K. and Salaminiou, E. (2001). Reduced positive future-thinking in depression: Cognitive and affective factors. *Cognition & Emotion*, 15(1):99–107.

MacLeod, C., Mathews, A., and Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, 95(1):15–20.

Maia, T. V. and McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: what participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences*, 101(45):16075–16080.

Maia, T. V. and McClelland, J. L. (2005). The somatic marker hypothesis: still many questions but no answers: Response to bechara et al. *Trends in Cognitive Sciences*, 9(4):162–164.

Maldonado, R., Saiardi, A., Valverde, O., Samad, T. A., Roques, B. P., and Borrelli, E. (1997). Absence of opiate rewarding effects in mice lacking dopamine D2 receptors. *Nature*, 388(6642):586–589.

Marler, P. and Evans, C. (1996). Bird calls: just emotional displays or something more? *Ibis*, 138(1):26–33.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* MIT Press, Massachusetts.

Mars, B., Heron, J., Gunnell, D., Martin, R. M., Thomas, K. H., and Kessler, D. (2017). Prevalence and patterns of antidepressant switching amongst primary care patients in the uk. *Journal of Psychopharmacology*, 31(5):553–560.

Massana, J. (1998). Reboxetine versus fluoxetine: an overview of efficacy and tolerability. *The Journal of Clinical Psychiatry*, 59(14):8–10.

Matell, M. S. and Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *Bioessays*, 22(1):94–103.

Matheson, S. M., Asher, L., and Bateson, M. (2008). Larger, enriched cages are associated with 'optimistic'response biases in captive european starlings (sturnus vulgaris). *Applied Animal Behaviour Science*, 109(2-4):374–383.

Mathews, A. and MacLeod, C. (1994). Cognitive approaches to emotion and emotional disorders. *Annual Review of Psychology*, 45(1):25–50.

Mathews, A., Richards, A., and Eysenck, M. (1989). Interpretation of homophones related to threat in anxiety states. *Journal of Abnormal Psychology*, 98(1):31–34.

McArthur, R. and Borsini, F. (2006). Animal models of depression in drug discovery: a historical perspective. *Pharmacology Biochemistry and Behavior*, 84(3):436–452.

McClure, S. M., Daw, N. D., and Montague, P. R. (2003). A computational substrate for incentive salience. *Trends in Neurosciences*, 26(8):423–428.

McDermott, L. M. and Ebmeier, K. P. (2009). A meta-analysis of depression severity and cognitive function. *Journal of Affective Disorders*, 119(1-3):1–8.

McKinnon, M. C., Yucel, K., Nazarov, A., and MacQueen, G. M. (2009). A meta-analysis examining clinical predictors of hippocampal volume in patients with major depressive disorder. *Journal of Psychiatry & Neuroscience*, 34(1):41–54.

Mendl, M., Burman, O. H., Parker, R. M., and Paul, E. S. (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science*, 118(3-4):161–181.

Mendl, M., Burman, O. H., and Paul, E. S. (2010). An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B: Biological Sciences*, 277(1696):2895–2904.

Mendl, M. and Paul, E. S. (2004). Consciousness, emotion and animal welfare: insights from cognitive science. *Animal Welfare*, 13(1):17–25.

Mendl, M. and Paul, E. S. (in prep). Animal affect and decision-making.

Mendl, M. and Paul, L. (2017). *Getting to the heart of animal welfare. The study of animal emotion.* Stichting Animales, Netherlands.

Metereau, E. and Dreher, J. (2012). Cerebral correlates of salient prediction error for different rewards and punishments. *Cerebral Cortex*, 23(2):477–487.

Meyer, G., Hauffa, B. P., Schedlowski, M., Pawlak, C., Stadler, M. A., and Exton, M. S. (2000). Casino gambling increases heart rate and salivary cortisol in regular gamblers. *Biological Psychiatry*, 48(9):948–953.

Meyniel, F., Goodwin, G. M., Deakin, J. W., Klinge, C., MacFadyen, C., Milligan, H., Mullings, E., Pessiglione, M., and Gaillard, R. (2016). A specific role for serotonin in overcoming effort cost. *Elife*, 5:e17282.

Mitchinson, B., Martin, C. J., Grant, R. A., and Prescott, T. J. (2007). Feedback control in active sensing: rat exploratory whisking is modulated by environmental contact. *Proceedings of the Royal Society B: Biological Sciences*, 274(1613):1035–1041.

Moberg, G. P. and Mench, J. (2000). *Biological response to stress: implications for animal welfare.* CABI, Wallingford.

Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *Annals of Internal Medicine*, 151(4):264–269.

Molendijk, M. L. and de Kloet, E. R. (2015). Immobility in the forced swim test is adaptive and does not reflect depression. *Psychoneuroendocrinology*, 62:389–391.

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of Neuroscience*, 16(5):1936–1947.

Montague, P. R., Dolan, R. J., Friston, K. J., and Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, 16(1):72–80.

Montgomery, K. (1955). The relation between fear induced by novel stimulation and exploratory drive. *Journal of Comparative and Physiological Psychology*, 48(4):254–260.

Mormède, P., Andanson, S., Aupérin, B., Beerda, B., Guémené, D., Malmkvist, J., Manteca, X., Manteuffel, G., Prunet, P., van Reenen, C. G., et al. (2007). Exploration of the hypothalamic-pituitary-adrenal function as a tool to evaluate animal welfare. *Physiology & behavior*, 92(3):317–339.

Morris, W. N. and Reilly, N. P. (1987). Toward the self-regulation of mood: Theory and research. *Motivation and Emotion*, 11(3):215–249.

Muris, P. and Van der Heiden, S. (2006). Anxiety, depression, and judgments about the probability of future negative and positive events in children. *Journal of Anxiety Disorders*, 20(2):252–261.

Murphy, J. G., Yurasek, A. M., Dennhardt, A. A., Skidmore, J. R., McDevitt-Murphy, M. E., MacKillop, J., and Martens, M. P. (2013). Symptoms of depression and PTSD are associated with elevated alcohol demand. *Drug and Alcohol Dependence*, 127(1-3):129–136.

Nakagawa, S., Johnson, P. C., and Schielzeth, H. (2017a). The coefficient of determination r2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, 14(134):20170213.

Nakagawa, S., Noble, D. W., Senior, A. M., and Lagisz, M. (2017b). Meta-evaluation of meta-analysis: ten appraisal questions for biologists. *BMC Biology*, 15(1):18.

Nakagawa, S. and Santos, E. S. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology*, 26(5):1253–1274.

Nakagawa, S. and Schielzeth, H. (2010). Repeatability for gaussian and non-gaussian data: a practical guide for biologists. *Biological Reviews*, 85(4):935–956.

Nakagawa, S. and Schielzeth, H. (2013). A general and simple method for obtaining r2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2):133–142.

National Research Council (1988). *Use of laboratory animals in biomedical and behavioral research*. National Academies Press, Washington.

Naudon, L. and Jay, T. (2005). Opposite behaviours in the forced swimming test are linked to differences in spatial working memory performances in the rat. *Neuroscience*, 130(2):285–293.

Nemeroff, C. B. (2003). The role of GABA in the pathophysiology and treatment of anxiety disorders. *Psychopharmacology Bulletin*, 37(4):133–146.

Nemeroff, C. B. and Owens, M. J. (2002). Treatment of mood disorders. *Nature Neuroscience*, 5:1068–1070.

Nesse, R. M. (2001). The smoke detector principle. *Annals of the New York Academy of Sciences*, 935(1):75–85.

Nestler, E. J., Barrot, M., DiLeone, R. J., Eisch, A. J., Gold, S. J., and Monteggia, L. M. (2002). Neurobiology of depression. *Neuron*, 34(1):13–25.

Nestler, E. J. and Hyman, S. E. (2010). Animal models of neuropsychiatric disorders. *Nature Neuroscience*, 13(10):1161–1169.

Nettle, D. (2008). An evolutionary model of low mood states. *Journal of Theoretical Biology*, pages 1–4.

Nettle, D. and Bateson, M. (2012). The evolutionary origins of mood and its disorders. *Current Biology*, 22(17):R712–R721.

Neumann, J. V. and Morgenstern, O. (1944). *Theory of games and economic behavior*, volume 60. Princeton University Press, Princeton.

Neville, V., Andrews, C., Nettle, D., and Bateson, M. (2017). Dissociating the effects of alternative early-life feeding schedules on the development of adult depression-like phenotypes. *Scientific Reports*, 7(1):14832.

Niv, Y., Daw, N. D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*, 191(3):507–520.

Niv, Y., Edlund, J. A., Dayan, P., and O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *Journal of Neuroscience*, 32(2):551–562.

Nonacs, P. (2001). State dependent behavior and the marginal value theorem. *Behavioral Ecology*, 12(1):71–83.

Nygren, T. E., Isen, A. M., Taylor, P. J., and Dulin, J. (1996). The influence of positive affect on the decision rule in risk situations: Focus on outcome (and especially avoidance of loss) rather than probability. *Organizational Behavior and Human Decision Processes*, 66(1):59–72.

Oatley, K. and Jenkins, J. M. (1992). Human emotions: Function and dysfunction. *Annual Review of Psychology*, 43(1):55–85.

Öhman, A., Flykt, A., and Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General*, 130(3):466–478.

Ore, O. (1960). Pascal and the invention of probability theory. *The American Mathematical Monthly*, 67(5):409–419.

Ottaviani, E. and Franceschi, C. (1996). The neuroimmunology of stress from invertebrates to man. *Progress in Neurobiology*, 48(4-5):421–440.

Otto, A. R., Fleming, S. M., and Glimcher, P. W. (2016). Unexpected but incidental positive outcomes predict real-world gambling. *Psychological Science*, 27(3):299–311.

Owens, M. J. and Nemeroff, C. B. (1994). Role of serotonin in the pathophysiology of depression: focus on the serotonin transporter. *Clinical Chemistry*, 40(2):288–295.

Padoa-Schioppa, C. and Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, 441(7090):223–226.

Padoa-Schioppa, C. and Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience*, 11(1):95–102.

Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *Behavioral and Brain Sciences*, 5(3):407–422.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford University Press, USA.

Papakostas, G. I. (2006). Dopaminergic-based pharmacotherapies for depression. *European Neuropsychopharmacology*, 16(6):391–402.

Parker, R. M. A., Paul, E. S., Burman, O. H. P., Browne, W. J., and Mendl, M. (2014). Housing conditions affect rat responses to two types of ambiguity in a reward-reward discrimination cognitive bias task. *Behavioural Brain Research*, 274:73–83.

Parr, T., Rees, G., and Friston, K. J. (2018). Computational neuropsychology and bayesian inference. *Frontiers in Human Neuroscience*, 12:61.

Parrott, A. C., Gibbs, A., Scholey, A. B., King, R., Owens, K., Swann, P., Ogden, E., and Stough, C. (2011). MDMA and methamphetamine: some paradoxical negative and positive mood changes in an acute dose laboratory study. *Psychopharmacology*, 215(3):527–536.

Patterson-Kane, E., Harper, D., and Hunt, M. (2001). The cage preferences of laboratory rats. *Laboratory Animals*, 35(1):74–79.

Paul, E. S., Cuthill, I., Kuroso, G., Norton, V., Woodgate, J., and Mendl, M. (2011). Mood and the speed of decisions about anticipated resources and hazards. *Evolution and Human Behavior*, 32(1):21–28.

Paul, E. S., Harding, E. J., and Mendl, M. (2005). Measuring emotional processes in animals: the utility of a cognitive approach. *Neuroscience & Biobehavioral Reviews*, 29(3):469–491.

Pauli, W. M., Larsen, T., Collette, S., Tyszka, J. M., Seymour, B., and O'Doherty, J. P. (2015). Distinct contributions of ventromedial and dorsolateral subregions of the human substantia nigra to appetitive and aversive learning. *Journal of Neuroscience*, 35(42):14220–14233.

Paulus, M. P. and Angela, J. Y. (2012). Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends in Cognitive Sciences*, 16(9):476–483.

Peciña, S. and Berridge, K. C. (2000). Opioid site in nucleus accumbens shell mediates eating and hedonic 'liking'for food: map based on microinjection fos plumes. *Brain Research*, 863(1-2):71–86.

Pellow, S., Chopin, P., File, S. E., and Briley, M. (1985). Validation of open: closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *Journal of Neuroscience Methods*, 14(3):149–167.

Perry, C. J., Baciadonna, L., and Chittka, L. (2016). Unexpected rewards induce dopamine-dependent positive emotion–like state changes in bumblebees. *Science*, 353(6307):1529–1531.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2018). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-131.1.

Platt, M. L. and Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, 400(6741):233.

Poirier, C., Bateson, M., Gualtieri, F., Armstrong, E. A., Laws, G. C., Boswell, T., and Smulders, T. V. (2019). Validation of hippocampal biomarkers of cumulative affective experience. *Neuroscience & Biobehavioral Reviews*, 101:113–121.

Poole, T. (1997). Happy animals make good science. *Laboratory Animals*, 31(2):116–124.

Porcelli, A. J. and Delgado, M. R. (2009). Acute stress modulates risk taking in financial decision making. *Psychological Science*, 20(3):278–283.

Porsolt, R., Le Pichon, M., and Jalfre, M. (1977). Depression: a new animal model sensitive to antidepressant treatments. *Nature*, 266(5604):730–732.

Posner, J., Russell, J. A., and Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(3):715–734.

Prescott, M. J. and Lidster, K. (2017). Improving quality of science through better animal welfare: the NC3Rs strategy. *Laboratory Animal*, 46(4):152–156.

Preston, S., Buchanan, T., Stansfield, R., and Bechara, A. (2007). Effects of anticipatory stress on decision making in a gambling task. *Behavioral Neuroscience*, 121(2):257–263.

Price, J. L. and Drevets, W. C. (2010). Neurocircuitry of mood disorders. *Neuropsychopharmacology*, 35(1):192.

Puterman, E., Lin, J., Krauss, J., Blackburn, E. H., and Epel, E. S. (2015). Determinants of telomere attrition over 1 year in healthy older women: stress and health behaviors matter. *Molecular Psychiatry*, 20(4):529–535.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Radenbach, C., Reiter, A. M., Engert, V., Sjoerds, Z., Villringer, A., Heinze, H.-J., Deserno, L., and Schlagenhauf, F. (2015). The interaction of acute and chronic stress impairs model-based behavioral control. *Psychoneuroendocrinology*, 53:268–280.

Rajan, R., Clement, J. P., and Bhalla, U. S. (2006). Rats smell in stereo. *Science*, 311(5761):666–670.

Ralph, C. and Tilbrook, A. (2016). Invited review: the usefulness of measuring glucocorticoids for assessing animal welfare. *Journal of Animal Science*, 94(2):457–470.

Ramboz, S., Oosting, R., Amara, D. A., Kung, H. F., Blier, P., Mendelsohn, M., Mann, J. J., Brunner, D., and Hen, R. (1998). Serotonin receptor 1a knockout: an animal model of anxiety-related disorder. *Proceedings of the National Academy of Sciences*, 95(24):14476–14481.

Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, 9(7):545–556.

Rao, R. P. and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87.

Reisenzein, R. (1994). Pleasure-arousal theory and the intensity of emotions. *Journal of Personality and Social Psychology*, 67(3):525–539.

Rescorla, R. A., Wagner, A. R., et al. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and Theory*, 2:64–99.

Reynolds, S. M. and Berridge, K. C. (2002). Positive and negative motivation in nucleus accumbens shell: bivalent rostrocaudal gradients for GABA-elicited eating, taste "liking"/"disliking" reactions, place preference/avoidance, and fear. *Journal of Neuroscience*, 22(16):7308–7320.

Reynolds, S. M. and Berridge, K. C. (2008). Emotional environments retune the valence of appetitive versus fearful functions in nucleus accumbens. *Nature Neuroscience*, 11(4):423–425.

Robinson, O. J., Overstreet, C., Charney, D. R., Vytal, K., and Grillon, C. (2013). Stress increases aversive prediction error signal in the ventral striatum. *Proceedings of the National Academy of Sciences*, 110(10):4129–4133.

Rock, P., Roiser, J., Riedel, W., and Blackwell, A. (2014). Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological Medicine*, 44(10):2029–2040.

Rodgers, R., Cao, B.-J., Dalvi, A., and Holmes, A. (1997). Animal models of anxiety: an ethological perspective. *Brazilian Journal of Medical and Biological Research*, 30:289–304.

Rolls, E. T. (2013). What are emotional states, and why do we have them? *Emotion Review*, 5(3):241–247.

Roseman, I. (2008). Motivations and emotivations: approach, avoidance, and other tendencies in motivated and emotional behavior. In *Handbook of approach and avoidance motivation*. Psychology Press, AJ Elliot, New York.

Rüedi-Bettschen, D., Zhang, W., Russig, H., Ferger, B., Weston, A., Pedersen, E.-M., Feldon, J., and Pryce, C. R. (2006). Early deprivation leads to altered behavioural, autonomic and endocrine responses to environmental challenge in adult Fischer rats. *European Journal of Neuroscience*, 24(10):2879–2893.

Ruhé, H. G., Mason, N. S., and Schene, A. H. (2007). Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: a meta-analysis of monoamine depletion studies. *Molecular Psychiatry*, 12(4):331–359.

Rupniak, N. (2003). Animal models of depression: challenges from a drug development perspective. *Behavioural Pharmacology*, 14(5):385–390.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110(1):145–172.

Russell, J. A., Lewicka, M., and Niit, T. (1989a). A cross-cultural study of a circumplex model of affect. *Journal of Personality and Social Psychology*, 57(5):848.

Russell, J. A., Weiss, A., and Mendelsohn, G. A. (1989b). Affect grid: a single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, 57(3):493–502.

Rutledge, R. B., Moutoussis, M., Smittenaar, P., Zeidman, P., Taylor, T., Hrynkiewicz, L., Lam, J., Skandali, N., Siegel, J. Z., Ousdal, O. T., et al. (2017). Association of neural and emotional impacts of reward prediction errors with major depression. *JAMA psychiatry*, 74(8):790–797.

Rutledge, R. B., Skandali, N., Dayan, P., and Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being. *Proceedings of the National Academy of Sciences*, 111(33):12252–12257.

Saito, Y., Yuki, S., Seki, Y., Kagawa, H., and Okanoya, K. (2016). Cognitive bias in rats evoked by ultrasonic vocalizations suggests emotional contagion. *Behavioural Processes*, 132:5–11.

Sanchez, A., Vazquez, C., Gomez, D., and Joormann, J. (2014). Gaze-fixation to happy faces predicts mood repair after a negative mood induction. *Emotion*, 14(1):85–94.

Sandem, A., Braastad, B., and Bøe, K. (2002). Eye white may indicate emotional state on a frustration–contentedness axis in dairy cows. *Applied Animal Behaviour Science*, 79(1):1–10.

Sanger, M. E., Doyle, R. E., Hinch, G. N., and Lee, C. (2011). Sheep exhibit a positive judgement bias and stress-induced hyperthermia following shearing. *Applied Animal Behaviour Science*, 131(3-4):94–103.

Sapolsky, R. M., Romero, L. M., and Munck, A. U. (2000). How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews*, 21(1):55–89.

Scheier, M. F., Carver, C. S., and Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the life orientation test. *Journal of Personality and Social Psychology*, 67(6):1063–1078.

Schein, M. W. and Orgain, H. (1953). A preliminary analysis of garbage as food for the Norway rat. *The American Journal of Tropical Medicine and Hygiene*, 2(6):1117–1130.

Scherer, K. R., Schorr, A., and Johnstone, T. (2001). *Appraisal processes in emotion: Theory, methods, research*. Oxford University Press.

Scholl, J. and Klein-Flügge, M. (2018). Understanding psychiatric disorder by capturing ecologically relevant features of learning and decision-making. *Behavioural Brain Research*, 355:56–75.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599.

Sclafani, A. and Ackroff, K. (2003). Reinforcement value of sucrose measured by progressive ratio operant licking in the rat. *Physiology & Behavior*, 79(4-5):663–670.

Segovia, G., Del Arco, A., De Blas, M., Garrido, P., and Mora, F. (2010). Environmental enrichment increases the in vivo extracellular concentration of dopamine in the nucleus accumbens: a microdialysis study. *Journal of Neural Transmission*, 117(10):1123–1130.

Seligman, M. E. (1972). Learned helplessness. *Annual review of medicine*, 23(1):407–412.

Senior, A. M., Grueber, C. E., Kamiya, T., Lagisz, M., O'Dwyer, K., Santos, E. S., and Nakagawa, S. (2016). Heterogeneity in ecological and evolutionary meta-analyses: its magnitude and implications. *Ecology*, 97(12):3293–3299.

Sescousse, G., Caldú, X., Segura, B., and Dreher, J.-C. (2013). Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, 37(4):681–696.

Shaham, Y. and Stewart, J. (1994). Exposure to mild stress enhances the reinforcing efficacy of intravenous heroin self-administration in rats. *Psychopharmacology*, 114(3):523–527.

Shaheed, C. A., Maher, C. G., Williams, K. A., Day, R., and McLachlan, A. J. (2016). Efficacy, tolerability, and dose-dependent effects of opioid analgesics for low back pain: a systematic review and meta-analysis. *JAMA Internal Medicine*, 176(7):958–968.

Sharp, M. E., Foerde, K., Daw, N. D., and Shohamy, D. (2015). Dopamine selectively remediates 'model-based'reward learning: a computational approach. *Brain*, 139(2):355–364.

Sherwin, C. (2004). The influences of standard laboratory cages on rodents and the validity of research data. *Animal Welfare*, 13(1):9–15.

Shettleworth, S. J. (2000). Modularity and the evolution of cognition. *The evolution of cognition, ed. CM Heyes & L. Huber*, pages 43–60.

Shipp, S., Adams, R. A., and Friston, K. J. (2013). Reflections on agranular architecture: predictive coding in the motor cortex. *Trends in Neurosciences*, 36(12):706–716.

Sierakowiak, A., Mattsson, A., Gómez-Galán, M., Feminía, T., Graae, L., Aski, S. N., Damberg, P., Lindskog, M., Brené, S., and Åberg, E. (2015). Hippocampal morphology in a rat model of depression: the effects of physical activity. *The Open Neuroimaging Journal*, 9:1–6.

Simmons, W. K., Burrows, K., Avery, J. A., Kerr, K. L., Bodurka, J., Savage, C. R., and Drevets, W. C. (2016). Depression-related increases and decreases in appetite: dissociable patterns of aberrant activity in reward and interoceptive neurocircuitry. *American Journal of Psychiatry*, 173(4):418–428.

Smallwood, R. D. and Sondik, E. J. (1973). The optimal control of partially observable markov processes over a finite horizon. *Operations Research*, 21(5):1071–1088.

Smith, H. R. and Porrino, L. J. (2008). The comparative distributions of the monoamine transporters in the rodent, monkey, and human amygdala. *Brain Structure and Function*, 213(1-2):73–91.

Smith, J. and Dickinson, A. (1998). The dopamine antagonist, pimozide, abolishes pavlovian-instrumental transfer. *Journal Psychopharmacology*, 12(S3):6.

Solinas, M., Chauvet, C., Thiriet, N., El Rawas, R., and Jaber, M. (2008). Reversal of cocaine addiction by environmental enrichment. *Proceedings of the National Academy of Sciences*, 105(44):17145–17150.

Spring, B., Pingitore, R., and McChargue, D. E. (2003). Reward value of cigarette smoking for comparably heavy smoking schizophrenic, depressed, and nonpatient smokers. *American Journal of Psychiatry*, 160(2):316–322.

Spruijt, B. M., Van den Bos, R., and Pijlman, F. T. (2001). A concept of welfare based on reward evaluating mechanisms in the brain: anticipatory behaviour as an indicator for the state of reward systems. *Applied Animal Behaviour Science*, 72(2):145–171.

Stankevicius, A., Huys, Q. J., Kalra, A., and Seriès, P. (2014). Optimism as a prior belief about the probability of future reward. *PLoS Computational Biology*, 10(5):e1003605.

Stearns, S. C. (2000). Daniel Bernoulli (1738): evolution and economics under risk. *Journal of Biosciences*, 25(3):221–228.

Steele, J. D., Meyer, M., and Ebmeier, K. (2004). Neural predictive error signal correlates with depressive illness severity in a game paradigm. *Neuroimage*, 23(1):269–280.

Stephan, K. E. and Mathys, C. (2014). Computational approaches to psychiatry. *Current Opinion in Neurobiology*, 25:85–92.

Steru, L., Chermat, R., Thierry, B., and Simon, P. (1985). The tail suspension test: a new method for screening antidepressants in mice. *Psychopharmacology*, 85(3):367–370.

Stewart, M., Webster, J., Schaefer, A., Cook, N., and Scott, S. (2005). Infrared thermography as a non-invasive tool to study animal welfare. *Animal Welfare*, 14(4):319–325.

Stone, L. A. and Nielson, K. A. (2001). Intact physiological response to arousal with impaired emotional recognition in alexithymia. *Psychotherapy and Psychosomatics*, 70(2):92–102.

Strunk, D. R., Lopez, H., and DeRubeis, R. J. (2006). Depressive symptoms are associated with unrealistic negative predictions of future life events. *Behaviour Research and Therapy*, 44(6):861–882.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, 3(1):9–44.

Sutton, R. S. (1990). Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In *Machine Learning Proceedings*, pages 216–224.

Sutton, R. S. and Barto, A. G. (1998). *Introduction to reinforcement learning*, volume 135. MIT press Cambridge.

Tamir, M. and Robinson, M. D. (2007). The happy spotlight: Positive mood and selective attention to rewarding information. *Personality and Social Psychology Bulletin*, 33(8):1124–1136.

Tanaka, M., Yoshida, M., Emoto, H., and Ishii, H. (2000). Noradrenaline systems in the hypothalamus, amygdala and locus coeruleus are involved in the provocation of anxiety: basic studies. *European Journal of Pharmacology*, 405(1-3):397–406.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychological Bulletin*, 110(1):67.

Thase, M. E. (1998). Effects of venlafaxine on blood pressure: a meta-analysis of original data from 3744 depressed patients. *The Journal of Clinical Psychiatry*, 59(10):502–508.

Thrun, S. B. (1992). *Efficient exploration in reinforcement learning*. Carnegie Mellon University, Pittsburgh.

Tobler, P. N., O'Doherty, J. P., Dolan, R. J., and Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *Journal of Neurophysiology*, 97(2):1621–1632.

Treadway, M. T., Bossaller, N. A., Shelton, R. C., and Zald, D. H. (2012). Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia. *Journal of Abnormal Psychology*, 121(3):553–558.

Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E., and Zald, D. H. (2009). Worth the 'eefrt'? the effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PloS One*, 4(8):e6598.

Trifilieff, P., Feng, B., Urizar, E., Winiger, V., Ward, R. D., Taylor, K. M., Martinez, D., Moore, H., Balsam, P. D., Simpson, E. H., et al. (2013). Increasing dopamine D2 receptor expression in the adult nucleus accumbens enhances motivation. *Molecular Psychiatry*, 18(9):1025–1033.

Trimmer, P. C., Ehlman, S. M., McNamara, J. M., and Sih, A. (2017). The erroneous signals of detection theory. *Proceedings of the Royal Society B: Biological Sciences*, 284(1865):20171852.

Trimmer, P. C., Houston, A. I., Marshall, J. A. R., Mendl, M. T., Paul, E. S., and McNamara, J. M. (2011). Decision-making under uncertainty: biases and Bayesians. *Animal Cognition*, 14(4):465–476.

Trimmer, P. C., Paul, E. S., Mendl, M. T., McNamara, J. M., and Houston, A. I. (2013). On the evolution and optimality of mood states. *Behavioral Sciences*, 3(3):501–521.

UK Home Office (2014). *Code of practice for the housing and care of animals bred, supplied or used for scientific purposes*. Controller of Her Majesty's Stationery Office London.

UK Home Office (2017). Annual statistics of scientific procedures on living animals, Great Britain 2016.

Uylings, H. B. and van Eden, C. G. (1991). Qualitative and quantitative comparison of the prefrontal cortex in rat and in primates, including humans. In *Progress in brain research*, volume 85, pages 31–62. Elsevier.

Valentin, V. V. and O'Doherty, J. P. (2009). Overlapping prediction errors in dorsal striatum during instrumental learning with juice and money reward in the human brain. *Journal of Neurophysiology*, 102(6):3384–3391.

Valvassori, S. S., Budni, J., Varela, R. B., and Quevedo, J. (2013). Contributions of animal models to the study of mood disorders. *Brazilian Journal of Psychiatry*, 35:S121–S131.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., and Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*, 45(2):576–594.

Van der Harst, J. E., Baars, A., and Spruijt, B. M. (2003). Standard housed rats are more sensitive to rewards than enriched housed rats as reflected by their anticipatory behaviour. *Behavioural Brain Research*, 142(1-2):151–156.

Vasquez, R. A. and Kacelnik, A. (2000). Foraging rate versus sociality in the starling sturnus vulgaris. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1439):157–164.

Veissier, I., Butterworth, A., Bock, B., and Roe, E. (2008). European approaches to ensure good animal welfare. *Applied Animal Behaviour Science*, 113(4):279–297.

Venter, J., Di Porzio, U., Robinson, D., Shreeve, S., Lai, J., Kerlavage, A., Fracek Jr, S., Lentes, K., and Fraser, C. (1988). Evolution of neurotransmitter receptor systems. *Progress in Neurobiology*, 30(2-3):105–169.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3):1–48.

Von Borell, E., Langbein, J., Després, G., Hansen, S., Leterrier, C., Marchant-Forde, J., Marchant-Forde, R., Minero, M., Mohr, E., Prunier, A., et al. (2007). Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals—a review. *Physiology & Behavior*, 92(3):293–316.

Von Neumann, J. and Morgenstern, O. (1953). *Theory of games and economic behavior*. Princeton University Press, New Jersey.

Wallace, D. J., Greenberg, D. S., Sawinski, J., Rulla, S., Notaro, G., and Kerr, J. N. (2013). Rats maintain an overhead binocular field at the expense of constant fusion. *Nature*, 498(7452):65–69.

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54(6):1063.

Watson, D., Wiese, D., Vaidya, J., and Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, 76(5):820–838.

Watts, J. M. and Stookey, J. M. (2000). Vocal behaviour in cattle: the animal's commentary on its biological processes and welfare. *Applied Animal Behaviour Science*, 67(1-2):15–33.

West, A. P. (1990). Neurobehavioral studies of forced swimming: the role of learning and memory in the forced swim test. *Progress in Neuro-psychopharmacology & Biological Psychiatry*, 14(6):863–877.

Whiteley, L. and Sahani, M. (2008). Implicit knowledge of visual uncertainty guides decisions with asymmetric outcomes. *Journal of Vision*, 8(3):1–15.

Wichmann, F. A. and Hill, N. J. (2001). The psychometric function: I. fitting, sampling, and goodness of fit. *Perception & Psychophysics*, 63(8):1293–1313.

Williams, J. M. G., Mathews, A., and MacLeod, C. (1996). The emotional stroop task and psychopathology. *Psychological bulletin*, 120(1):3.

Wilson, R. and Collins, A. (2019). Ten simple rules for the computational modeling of behavioral data. *PsyArXiv*.

Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., et al. (2017). Drugbank 5.0: a major update to the drugbank database for 2018. *Nucleic Acids Research*, 46(D1):D1074–D1082.

World Health Organization (2017). Depression and other common mental disorders: global health estimates, 2017.

Wright, W. F. and Bower, G. H. (1992). Mood effects on subjective probability assessment. *Organizational Behavior and Human Decision Processes*, 52(2):276–291.

Wulfert, E., Roland, B. D., Hartley, J., Wang, N., and Franco, C. (2005). Heart rate arousal and excitement in gambling: winners versus losers. *Psychology of Addictive Behaviors*, 19(3):311–316.

Wunderlich, K., Smittenaar, P., and Dolan, R. J. (2012). Dopamine enhances model-based over model-free choice behavior. *Neuron*, 75(3):418–424.

Yeates, J. W. and Main, D. C. (2008). Assessment of positive welfare: a review. *The Veterinary Journal*, 175(3):293–300.

Yuen, K. S. and Lee, T. M. (2003). Could mood state affect risk-taking decisions? *Journal of Affective Disorders*, 75(1):11–18.

Zhang, J., Berridge, K. C., Tindell, A. J., Smith, K. S., and Aldridge, J. W. (2009). A neural computational model of incentive salience. *PLoS Computational Biology*, 5(7):e1000437.

Zhu, J., Apparsundaram, S., Bardo, M. T., and Dwoskin, L. P. (2005). Environmental enrichment decreases cell surface expression of the dopamine transporter in rat medial prefrontal cortex. *Journal of Neurochemistry*, 93(6):1434–1443.

Zimmerman, P. and Koene, P. (1998). The effect of frustrative nonreward on vocalisations and behaviour in the laying hen, gallus gallus domesticus. *Behavioural Processes*, 44(1):73–79.

Zimmerman, P., Lundberg, A., Keeling, L., and Koene, P. (2003). The effect of an audience on the gakel-call and other frustration behaviours in the laying hen (gallus gallus domesticus). *Animal Welfare*, 12(3):315–326.